Supplemental Information

Inventory of Supplementary Materials:

Figure S1, related to Figure 1.

The ARD of AKR2A exhibits chloroplast binding.

Figure S2, related to Figure 3.

Binding of His:AKR2A to wild-type, *mgd1* and *pgp1-1* chloroplasts at various conditions *in vitro*.

Figure S3, related to Figure 4.

Chloroplast targeting of GFP:AtToc34 is impaired in *mgd1* and *pgp1-1* mutant protoplasts.

Figure S4, related to Figure 6.

Amino acid residues critical for lipid binding are highly conserved in AKR2 homologs.

Figure S5, related to Figure 7.

The AKR2 gene phylogeny.

Table S1, related to Figure 5.

Statistics of data collection and refinement

Supplemental Text, including:

Supplemental Figure Legends

Extended Experimental Procedures

Supplemental References

SUPPLEMENTAL FIGURE LEGENDS

Figure S1, related to Figure 1. The ARD of AKR2A exhibits chloroplast binding.

(A) Various AKR2A deletion mutants used in this study are presented schematically. (B and C) A minimal AKR2A domain for chloroplast binding. Full-length and various deletion mutants of AKR2A were used in chloroplast binding experiments *in vitro*. (B) Western blot analysis of chloroplast binding of full-length and deletion mutants of AKR2A. Cab, loading control. (C) Quantification of chloroplast binding of various AKR2A deletion mutants. Mean ± standard deviation (SD) are shown (n = 3). The asterisks indicate a significant difference from the corresponding control experiment by Student's *t*-test (*$P < 0.05$; **$P < 0.01$; ***$P < 0.001$). (D) Effect of sHsp17.8 on AKR2A binding to chloroplasts. Purified His:sHsp17.8 were incubated with chloroplasts and His:sHsp17.8-containing chloroplasts were used for AKR2A binding with or without trypsin treatment. In the case of the trypsin treatment, chloroplasts were repurified before binding experiments. The amount of His:AKR2A bound to chloroplasts was determined by Western blot analysis using anti-His, anti-AtToc75 and anti-Toc159 antibodies. The AtToc75 and AtToc159 levels were determined as internal controls for the trypsin treatment. Notice that under our tryptic digest condition of His:sHsp17.8, two endogenous COM proteins AtToc75 and AtToc159 were quantitatively removed (Figure S1D).

Figure S2, related to Figure 3. Binding of His:AKR2A to wild-type, *mgd1* and *pgp1-1* chloroplasts at various conditions *in vitro*.

(A and B) The effect of yPGC1 treatment on AKR2A binding to chloroplasts. (A) His:AKR2A was incubated with chloroplasts that had been treated with recombinant His:yPGC1 at varying incubation times and the amount of His:AKR2A copurified with chloroplasts was determined by Western blot analysis using

anti-His antibody. (B) Quantification of His:AKR2A binding to His:yPGC1-treated chloroplasts. Mean ± SD are shown (n = 3). (C and D) The effect of MGDG and PG supplementation to *mgd1* and *pgp1-1* chloroplasts on AKR2A binding, respectively. (C) Chloroplast envelope membranes from *mgd1* chloroplasts were supplemented with MGDG at two different concentrations and used to prepare liposomes by sonication. The liposomes of chloroplast envelope membranes was incubated with His:AKR2A in the presence or absence of duramycin and the amount of His:AKR2A pelleted with the liposomes was determined by Western blot analysis using anti-His antibody. (D) Chloroplast envelope membranes from *pgp1-1* chloroplasts were supplemented with PG at two different concentrations and used to prepare liposomes by sonication. The liposomes were incubated with His:AKR2A in the presence or absence of daptomycin and the amount of His:AKR2A pelleted with the liposomes was determined by Western blot analysis using anti-His antibody. (E and F) As control experiments to (C) and (D), chloroplast envelope membranes from *mgd1* or *pgp1-1* chloroplasts were supplemented with the indicated lipids and used to prepare liposomes by sonication. Liposomes of chloroplast envelope membranes were incubated with His:AKR2A and the amount of His:AKR2A pelleted with liposomes was determined by Western blot analysis using anti-His antibody. (G) Quantification of His:AKR2A binding to PS or PI supplemented liposomes. In the case of PS, the amount of AKR2A bound to PS supplemented liposomes was slightly increased. Accordingly, the data were quantified to access the increase. Mean ± SD are shown (n = 3). The asterisks indicate a significant difference from the corresponding control experiment by Student's *t*-test (*$P < 0.05$; **$P < 0.01$; ***$P < 0.001$).

Figure S3, related to Figure 4. Chloroplast targeting of GFP:AtToc34 is impaired in *mgd1* and *pgp1-1* mutant protoplasts.

(A and B) *GFP:AtToc34* was cotransformed into the WT, *mgd1* or *pgp1-1* protoplasts with *mRFP*. To

determine the targeting efficiency of GFP:AtToc34, protein extracts from the transformed protoplasts (To, total) or from chloroplast fractions of transformed protoplasts (CH, chloroplast) were analyzed by western blotting using anti-GFP and anti-RFP antibodies (A). RbcL was used as a loading control. NT, non-transformed; To, total protoplast extracts; CH, chloroplast. (B) The band intensity of chloroplast fractions, presented as relative values to that of the WT. Mean ± SD are shown (n = 3). (C) The expression level of *AtToc75* and *AtToc34* in the WT and *mgd1* mutant plants. Total RNA from 2-week-old WT and *mgd1* mutant plants was subjected to real-time quantitative RT-PCR analysis using gene-specific primers of the indicated genes. *ACT2* was used as an internal control. Mean ± SD are shown (n = 3). The asterisks indicate a significant difference from the corresponding control experiment by Student's *t*-test (*$P < 0.05$; **$P < 0.01$; ***$P < 0.001$).

Figure S4, related to Figure 6. Amino acid residues critical for lipid binding are highly conserved in AKR2 homologs.

An alignment of key residues of ARDs responsible for lipid binding (H223, E246, F257, Y261, Y290, Y294, and R296) (highlighted), which are highly conserved in land plants and some green algae but not in cyanobacteria.

Figure S5, related to Figure 7. The AKR2 gene phylogeny.

Maximum likelihood (ML) tree of AKR2 and homologous sequences from cyanobacteria, green algae and land plants. The full length (857 amino acid) versions of the top 40 sequences were selected from a BLASTp search (e-value = 1e-10) of the NCBI RefSeq database (version 55). The ML tree was constructed by using RAxML under the LG+F+G protein model; numbers near the node indicate the ML bootstrap support value from 1,000 replications.

Table S1, related to Figure 5. Statistics of data collection and refinement

Data collection	
Space group $P2_1$	
Cell dimensions a, b, c (Å)	68.33 59.38 84.85
Cell angles α, β, γ (°)	90 107.33 90
Resolution (Å)	50 - 2.3
Reflections	
Total (N)	765,327
Unique (N)	33,241
Average (I/σ) (last shell; 2.37-2.3)[a]	27.5(3.4)
R_{sym} (%) (last shell; 2.37-2.3)[a]	8.5 (51.8)
Completeness (%) (last shell; 2.37-2.3)[a]	98.7 (98.7)
Refinement	
Resolution range (Å)	20-2.3
Number of reflections	28746
Number of protein atoms	4500
Number of water atoms	69
$R_{working}$ (%)	20.65
R_{free} (%)	22.41
r.m.s.d from ideality	
Bond length (Å)	0.008
Bond angles (°)	1.398

$R_{sym}=\Sigma_h\Sigma_i|I_{h,i}-I_h|/\Sigma_h\Sigma_iI_{h,i}$, where I_h is the mean intensity of i observations of symmetry-related reflections of h.

$R=\Sigma|F_{obs}-F_{calc}|/\Sigma F_{obs}$, where $F_{obs}=F_p$, and F_{calc} is the calculated protein structure factor from the atomic model (R_{free} was calculated with 5% of the reflections). R.m.s.d. values in bond lengths and angles are deviations from ideal values, and r.m.s.d. in B-factors is calculated between bonded atoms.

SUPPLEMENTAL EXPERIMENTAL PROCEDURES

Construction of plasmids

The constructs of His-tagged recombinant proteins of various ARD alanine substitution mutants and yPGC1 were generated by a PCR approach. yPGC1 has a single transmembrane domain at the end of the C-terminus. *yPGC1* without the TMD was amplified by PCR using a gene-specific primer set. The sequences of primers used in this study were listed in supplementary experimental procedures, primers. The primers used were as follows: ARD-H223A-F and ARD-His-R for *His:ARD(H223A)*; ARD-His-F and ARD-K338E-R for *His:ARD(K338E)*; and yPGC1-F and yPGC1-R for *His:yPGC1*. The other ARD alanine substitution mutants were prepared by two sequential overlapping PCRs. The primers were as follows: ARD-His-F/ARD(E246A)-R and ARD(E246A)-F/ARD-His-R for the 1st PCR of *His:ARD(E246A)*; ARD-His-F/ARD(Y294A)-R and ARD(Y294A)-F/ARD-His-R for the 1st PCR of *His:ARD(Y294A)*; ARD-His-F/ARD(R296A)-R and ARD(R296A)-F/ARD-His-R for the 1st PCR of *His:ARD(R296A)*; ARD-His-F/ARD(N314A)-R and ARD(N314A)-F/ARD-His-R for the 1st PCR of *His:ARD(N314A)*; ARD-His-F/ARD(L315A)-R and ARD(L315A)-F/ARD-His-R for the 1st PCR of *His:ARD(L315A)*; and ARD-His-F and ARD-His-R for the 2nd PCR of *His:ARD(E246A)*, *His:ARD(Y294A)*, *His:ARD(R296A)*, *His:ARD(N314A)*, and *His:ARD(L315A)*. PCR products were ligated into pRSET-A for His-tagging. The primers used for generation of HA-tagged AKR2A deletion mutants were as follows: AKR2AΔ(PEST)-F and AKR2A-HA-R for *AKR2AΔ(PEST):HA*; AKR2AΔ(PEST+C1)-F and AKR2A-HA-R for *AKR2AΔ(PEST+C1):HA*; ARD-HA-F and AKR2A-HA-R for *ARD:HA*. To generate HA-tagged AKR2A(Y294A), PCR was performed using AKR2A-HA-F/ARD(Y294)-R and ARD(Y294A)-F/AKR2A-HA-R for the 1st PCR of *AKR2A(Y294A):HA*; and AKR2A-HA-F and AKR2A-HA-R for the 2nd PCR of *AKR2A(Y294A):HA*. The PCR products were ligated into -HA tagging vectors. To generate the *mCherry:annexin V* construct, *annexin V* was amplified by PCR. The PCR products were inserted into the

downstream of *mCherry* in *pUC-mCherry*.

AKR2A binding to chloroplasts and liposomes of chloroplast envelope membranes *in vitro*

Intact chloroplasts were isolated from protoplasts using a Percoll gradient as described previously (Li and Chen, 1996; Tu and Li, 2000). Purified His-tagged recombinant proteins (2 μg) were incubated with intact chloroplasts (equivalent to 20 μg chlorophyll mL^{-1}) in 1 ml binding buffer [50 mM Hepes-KOH, pH 7.6, 3 mM $MgCl_2$, 330 mM sorbitol, 100 mM NaCl, 1 mM DTT, 1 x complete protease inhibitor cocktail (Roche)] on ice for 30 min. Chloroplasts were pelleted by centrifugation (1,500 x g) for 5 min at 4°C. The pellet fraction was subjected to Western blot analysis using an anti-His antibody. Trypsin pretreatment of WT chloroplasts or His:sHsp17.8-bound chloroplasts was performed by incubating chloroplasts (20 μg of chlorophyll mL^{-1}) with trypsin at 22°C in the dark for 30 min (Tu and Li, 2000). The trypsin treatment was terminated by adding complete protease inhibitor cocktail. Chloroplasts were repurified by using Percoll cushion, washed in binding buffer, and used for an in vitro binding assay.

For duramycin treatment, purified chloroplasts or liposomes of chloroplast envelope membranes were incubated with duramycin (50 μM) in binding buffer on ice for 30 min in the dark (Tu and Li, 2000). For daptomycin treatment, intact chloroplasts or liposomes of chloroplast envelope membranes were incubated with daptomycin (1 to 10 μg/ml) in binding buffer with calcium (5 mM) on ice for 20 min in the dark (Muraih et al., 2011). For yPGC1 treatment, purified chloroplasts were incubated with purified His:yPGC1 (1 to 10 μg/ml) in binding buffer with 5 mM $CaCl_2$ at room temperature for 0 to 20 min in the dark. The His:yPGC1 treatment was terminated by the rapid cooling to 0°C. Intact chloroplasts were repurified using a Percoll cushion, followed by washing with binding buffer.

For preparation of chloroplast envelope membranes, intact chloroplasts were lysed in a hypotonic solution (10 mM Tricine-NaOH, pH 7.8) and the chloroplast envelope membranes were separated from the

chloroplast lysates by ultracentrifugation using a sucrose step gradient as described previously (Douce and Joyard, 1982). To supplement MGDG or PG to envelope membranes of *mgd1* or *pgp1-1* chloroplasts, respectively, MGDG or PG in chloroform/methanol was placed in siliconized tubes and the solvent was evaporated with N_2 gas. Subsequently, vesicles (200 μl) of chloroplast envelope membranes (0.5 mg protein/ml) were added to the tubes and liposomes were prepared from the lipid-supplemented chloroplast envelope membranes by sonication. For AKR2A binding to the liposomes of chloroplast envelope membranes, purified His:AKR2A (0.5 μg) was added and incubated at 22°C for 10 min. The liposomes were pelleted by ultracentrifugation at 100,000 x g for 1 h at 22°C and subjected to Western blot analysis using an anti-His antibody.

RNA extraction and real-time quantitative RT-PCR analysis

For quantitative RT-PCR (qRT-PCR) analysis, total RNA was extracted from 2-weeks-old seedlings using the RNAqueous kit (Ambion). TURBO DNase (Ambion) was used to remove any DNA contamination from the RNA samples. cDNA was generated from the extracted RNA using the High Capacity cDNA Reverse Transcription Kit (AB) and was subjected to qRT-PCR analysis of *AtToc75* and AtToc34 transcript levels. qRT-PCR was performed using a real-time PCR system (StepOne; Applied Biosystems, USA) using the SYBR Green Kit (Applied Biosystems, USA). *ACT2* was used as an internal control for qRT-PCR (Xu et al. 2012). The primer sets used were as follows: Toc75-III-RT-F and Toc75-III-RT-R for *AtToc75*; Toc34-RT-F and Toc34-RT-R for *AtToc34*, and ACT2-RT-F and ACT2-RT-R for *ACT2*.

Sequence alignment and generation of a maximum likelihood tree

Arabidopsis AKR2A sequence was used as query to search NCBI RefSeq database (version 55) using BLASTp (-e = 1e-10). Among the 392 matched sequences, the top 100 hits were retrieved for phylogenetic

analysis. In addition, to further examine the relationship of the ARDs between AKR2A and cyanobacterial ARD-harboring proteins, four cyanobacterial sequences (*Cyanothece* sp. PCC7424_GI.218439067, *Synechococcus* sp. JA-3-3Ab_GI.86607279, *Synechococcus* sp. JA-2-3Ba2-13_GI.86609865, and *Cyanothece* sp. PCC7822_GI.307150916), which ranked 222nd, 247th, 295th and 386th in the BLASTp search, respectively, were included in the phylogenetic analysis. MAFFT (Katoh et al., 2005) software was used for alignment. The poorly aligned regions were removed using Gblocks (-b4=5, -b5=h) (Talavera and Castresana, 2007). Identical sequences from same taxa were removed. A maximum likelihood tree was constructed by using RAxML v.7.2.8 (Stamatakis, 2006) under L40 (Le et al., 2008)+F+G model (F, empirical amino acid frequencies; G, gamma distributed rate heterogeneity). We used '-f a' option for the best tree likelihood and rapid bootstrap analyses with '-# 1000' (no. bootstrap replications), '-i' (automatically optimized SPR branch rearrangement) and '-c' (25 distinct rate categories) options.

Primers

The sequences of primers used in this study are below:

Primer name	Sequence (5' $\rightarrow$ 3')
ARD-His-F	CCCTCGAGGTAGCAGAAGAAGGTGAA
ARD-His-R	GGAATTCTCAAAGGAAAGCATCCTTCTC
AKR2A-HA-F	GCTCTAGAATGGCTTCCAATTCGGAG
AKR2A-HA-R	TCCCCCGGGAAGGAAAGCATCCTTCTC
ARD-HA-F	GCTCTAGAATGGTAGCAGAAGAAGGTGAA
AKR2AΔ(PEST)-F	GCTCTAGAATGGCTATGGCCGGCTTGAATTTC
AKR2AΔ(PEST+C1)-F	GCTCTAGAATGAACTTTGATCCTCAACAG
ARD-H223A-F	CCCTCGAGGTAGCAGAAGAAGGTGAAGAAGAAGAGTCTATTGTTGCCCAA

ARD-E246A-F AACAAAGATGCAGAAGATTCTGAAGGA

ARD-E246A-R AGAATCTTCTGCATCTTTGTTACCACC

ARD-Y294A-F TGCTGCTGGTGCCGGGAGGAAAGAGTGT

ARD-Y294A-R TTTCCTCCCGGCACCAGCAGCATAATG

ARD-R296A-F TGGTTACGGGGCGAAAGAGTGTGTAAGC

ARD-R296A-R ACACTCTTTCGCCCCGTAACCAGCAGC

ARD-N314A-F ACTCTGCAAGCCCTAGACGAGAAGACG

ARD-N314A-R CTCGTCTAGGGCTTGCAGAGTGACTGC

ARD-L315A-F TCTGCAAAACGCAGACGAGAAGACGCCA

ARD-L315A-R CTTCTCGTCTGCGTTTTGCAGAGTGAC

ARD-K338E-R CGGAATTCTCAAAGGAAAGCATCCTCCTCA

PGC1-F CGGGATCCATGGTTGAAATTGTGGGCCAC

PGC1-R CGGAATTCTCATTTGGAATATAGAAGCGT

Annexin V-F CGGGATCCATGGCACAGGTTCTCAGAGGC

Annexin V-R GGGGTACTTAGTCATCTTCTCCACAGAG

Toc75-III-RT-F CTTACCTAACCGAGATACAGAGA

Toc75-III-RT-R GAGGCTACGGCTAGAGGTTTA

Toc34-RT-F TTTAATGGCAACAAGGCGATT

Toc34-RT-R TGGGTTTGGCCCTTCGA

ACT2-RT-F TATGAATTACCCGATGGGCAAG

ACT2-RT-R TGGAACAAGACTTCTGGGCAT

SUPPLEMENTAL REFERENCES

Douce, R., and Joyard, J. (1982). Purification of the chloroplast envelope. In methods in chloroplast molecular biology. Edelman, M., Hallick, R.B., and Chua, N.H., editors, Elsevier/North Holland Biomedical Press, Amsterdam. 351-358.

Katoh, K. Kuma, K. Toh, H., and Miyata, T. (2005). MAFFT version 5: improvement in accuracy of multiple sequence alignment. Nucleic Acids Res. *33*, 511-518.

Le, S.Q., Lartillot, N., and Gascuel, O. (2008). Phylogenetic mixture models for proteins. Philos Trans R Soc B Biol. Sci. *363*: 3965-3976.

Li, H.M., and Chen, L.J. (1996). Protein targeting and integration signal for the chloroplastic outer envelope membrane. Plant Cell *8*, 2117-2126.

Stamatakis, A. (2006). RAxML-VI-HPC: maximum likelihood-based phylogenetic analyses with thousands of taxa and mixed models. Bioinformatics *22*, 2688-2690.

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

Tu, S.L., and Li, H.M. (2000). Insertion of OEP14 into the outer envelope membrane is mediated by proteinaceous components of chloroplasts. Plant Cell *12*, 1951-1960.

Xu, Z.Y., Lee, K.H., Dong, T., Jeong, J.C., Jin, J.B., Kanno, Y., Kim, D.H., Kim, S.Y., Seo, M., Bressan, R.A., Yun, D.J., and Hwang, I. (2012). A vacuolar b-glucosidase homolog that possesses glucose-conjugated abscisic acid hydrolyzing activity plays an important role in osmotic stress responses in Arabidopsis. Plant Cell *24*, 2184-2199.