Evidence for a protein transported through the secretory pathway en route to the higher plant chloroplast

Arsenio Villarejo[1], Stefan Burén[1], Susanne Larsson[1], Annabelle Déjardin[1†], Magnus Monné[2], Charlotta Rudhe[2], Jan Karlsson[1], Stefan Jansson[1], Patrice Lerouge[3], Norbert Rolland[4], Gunnar von Heijne[2], Markus Grebe[5], Laszlo Bako[1,6] and Göran Samuelsson[1*]

[1]Umeå Plant Science Centre, Department of Plant Physiology, Umeå University, S-90187 Umeå, Sweden

[2]Department of Biochemistry and Biophysics, Arrhenius Laboratory, Stockholm University, S-10691, Stockholm, Sweden

[3]UMR-CNRS 6037, IFRMP 23, UFR des Sciences, University of Rouen F-76821 Mont Saint Aignan, France

[4]Laboratoire de Physiologie Cellulaire Végétale, UMR-5168 CNRS/INRA/Université Joseph Fourier/CEA Grenoble, 17 rue des Martyrs, F-38054 Grenoble cedex 9, France

[5]Umeå Plant Science Centre, Department of Forest Genetics and Plant Physiology, Swedish University for Agricultural Sciences, S-90183 Umeå, Sweden

[6]Institute of Plant Biology, Biological Research Center of the Hungarian Academy of Sciences, H-6701 Szeged, Hungary

[†]Present address: Unité d'Amélioration, de Génétique et de Physiologie Forestières, INRA, BP 20619 Ardon, F-45166 Olivet Cedex, France

In contrast to animal and fungal cells, green plant cells contain one or multiple chloroplasts, the organelle(s) where photosynthetic reactions take place. Chloroplasts are believed to have originated from an endosymbiotic event and contain DNA coding for some of their proteins. Most chloroplast proteins are encoded by the nuclear genome and imported with the help of sorting signals that are intrinsic parts of the polypeptides. Here, we show that a chloroplast-located protein in higher plants takes an alternative route through the secretory pathway, and becomes *N*-glycosylated before entering the chloroplast.

In eukaryotic cells, sorting of newly synthesized proteins to sub-cellular compartments is a fundamental process in all organisms. Chloroplasts are believed to have originated from an endosymbiotic event in which a prokaryotic cell with its own genome complement was engulfed by an eukaryotic host[1]. During evolution, most endosymbiont genes were transferred from the progenitor organelle to the host cell nucleus[1,2]. Accordingly, the vast majority of chloroplast proteins is encoded by the nuclear genome and synthesized in precursor form on cytosolic ribosomes[3,4]. Chloroplast precursor proteins contain cleavable transit peptides, which direct them to the chloroplast in an organelle-specific way[4,5]. Translocation into chloroplasts occurs post-translationally and involves binding of precursor polypeptides to the Toc/Tic apparatus in the chloroplast envelope[4,5]. Recent studies of the chloroplast proteome have revealed the occurrence of many proteins without predicted transit peptides[6,7], indicating that targeting to chloroplasts may be more complex than first expected[8].

In our studies on chloroplast-localized proteins, we identified an *Arabidopsis* EST (Z18493) coding for an α–carbonic anhydrase (α-CA). The respective cDNA,

which we denoted *CAH1*, contains an 1046 bp open reading frame coding for a polypeptide of 284 amino acids. The CAH1 protein shares approximately 30 % amino acid sequence identity and 45% similarity with other α-CA sequences. It contains all 15 conserved catalytic and zinc binding residues typical for active α-CAs[9] (Fig. 1a, green and blue residues). Like secreted CAs, CAH1 has an extended amino-terminal sequence compared with the human CAs I, II and III, (Fig. 1a, italics). TargetP analyses [10,11] predict that this sequence targets the protein to the endoplasmic reticulum (ER) and locate a potential signal peptidase cleavage site between amino acids 24 and 25 (ADA|Q, Fig. 1a, arrow). Stop codons in all three reading frames are found upstream of the putative initiator Met, ruling out the existence of an additional targeting sequence. To localize CAH1 within the cell, we performed immunolocalization analysis in *Arabidopsis* leaves. Unexpectedly, the results indicated that CAH1 is located exclusively in the chloroplast stroma (Fig. 1b, c, see Supplementary Information, Fig. S1a). To confirm this localization pattern we analysed the presence of CAH1 in highly pure chloroplast and sub-plastid preparations (see Supplementary Information, Fig. S1b). CAH1 was enriched in intact chloroplasts and the stroma fraction, similarly to the enrichment of the large subunit of Rubisco (see band above 47 kDa), when compared to total cell extracts (Fig. 1d). To exclude the possibility of a contamination of the plastid fraction, we performed additional experiments. First, we tested the thermolysin susceptibility of CAH1. CAH1 was completely resistant to thermolysin treatment of intact chloroplasts, but was found to be susceptible after lysis of the chloroplasts (see Supplementary Information, Fig. S1c). All these results support a stroma localization of CAH1. As an independent verification of the CAH1 localization pattern, a translational fusion of green fluorescent protein (GFP) with the C-terminus of *Arabidopsis* CAH1 was

transiently expressed in *Arabidopsis* cells. The CAH1-GFP fusion protein was targeted to the chloroplasts in these cells (Fig. 1e). Non-fused GFP protein (negative control) was distributed uniformly throughout the cytosol and nucleus (Fig. 1f), whereas GFP fused to the transit peptide sequence of RbcS was targeted to chloroplasts (Fig. 1g). Thus, it appears that the CAH1 sequence information is sufficient for chloroplast targeting of the fusion protein *in vivo*.

We next determined the functionality of the predicted CAH1 signal peptide (SP). The N-terminal 40 amino acid residues of CAH1, containing the predicted ER SP, were fused to GFP harbouring an ER retention signal (KDEL) at the C-terminus. When expressed in plant cells, this fusion protein was retained in the ER (Fig. 2a), as shown by co-localization with a selective marker for the ER (ER-Tracker, Molecular Probes, Or, USA), indicating that the CAH1 SP is functional and sufficient for targeting the protein to the ER. This was confirmed by N-terminal sequencing of the gene product of this construct. The cleavage site was located between ADAQ/T, almost exactly at the site predicted by TargetP. In addition, when the full-length protein was fused to GFP containing a C-terminal KDEL sequence, the fusion protein was also retained in the ER. The GFP signal precisely overlapped with that of the ER tracker (Fig. 2b) as well as ER-targeted DsRed2-ER (see Supplemetary Information, Fig. S1d). Secondly, we performed uptake studies both with isolated chloroplasts and ER-derived dog pancreas microsomes[12]. Intriguingly, intact pea chloroplasts were not able to take up or process the full-length CAH1 precursor or CAH1 lacking the SP (see Supplemetary Information, Fig. S1e), suggesting that translocation of CAH1 across the envelope membranes may not take place through the Toc/Tic apparatus. In contrast, efficient uptake, SP removal, and glycosylation were observed in microsome preparations (Fig. 2c). The SP proved to be required for uptake of the protein into the

microsomes, since a truncated CAH1 form, lacking SP was not taken up into the ER as evidenced by lack of glycosylation and sensitivity to externally added proteinase K (Fig. 2c, lanes 4, 8-10). With full-length CAH1, the SP is cleaved off after import, leading to a shift in mobility (Fig. 2c, lanes 3, 4). These findings demonstrate *in vitro* targeting of CAH1 to microsomes, but not chloroplasts, via its N-terminal SP.

The CAH1 protein has five predicted acceptor sites for *N*-linked glycosylation (Fig. 1a, underlined triplets), and products with relative molecular masses of 38, 41 and 44 kDa are observed in addition to the CAH1 precursor and the unglycosylated, signal-peptidase processed form of the protein (Fig. 2c, lane 2). Addition of a competitive glycosylation inhibitor peptide prevents the appearance of the high molecular weight products (Fig. 2c, lane 3). The glycosylated and the unglycosylated signal-peptidase processed forms of the protein are resistant to externally added proteinase K (Fig. 2c, lane 7), and are thus located in the lumen of microsomes. These findings suggest that CAH1 is not only taken up into the ER but also glycosylated prior to being targeted to the chloroplast.

We addressed potential glycosylation of CAH1 in an independent experiment. CAH1-GFP and CAH1-GFP-KDEL proteins from transformed *Arabidopsis* protoplasts were immuno-precipitated with either anti-GFP or anti-CAH1 antibodies, and immunoblots were probed with the respective antibody that had not been used for immuno-precipitation (see Supplemetary Information, Fig. S2a, b). As expected, both antibodies precipitated and detected a protein of the same size demonstrating specificity of the CAH1 antibody. We then examined glycosylation of CAH1-GFP by probing the blots of these immunocomplexes with $\alpha(1,3)$-fucose antibodies. Only the CAH1-GFP fusion was detected, while the construct carrying the KDEL ER retention signal did not show fucosylation (Fig. 3a). This result suggests that CAH1-GFP

leaves the ER for the Golgi, since this is the organelle where the fucose transferase is located[13], and ER retention of CAH1-GFP prevents this fucosylation.

We next examined whether glycosylated CAH1 accumulates in the chloroplast and if this could explain the discrepancy between the apparent molecular mass of the mature protein isolated from chloroplasts (~38 kDa) and the non-glycosylated product obtained *in vitro* (30 kDa). To this end, a highly purified stroma fraction enriched in CAH1 was obtained and analyzed using antibodies specific for ß(1,2)-xylose and α(1,3)-fucose epitopes[14], (see Supplemetary Information, Fig. S2c-f) known to be typical for complex-type *N*-glycans of plants. A number of proteins were immunodecorated with the antibodies indicating that several stroma proteins are *N*-glycosylated. Prominent spots recognized by the β(1,2)-xylose and α(1,3)-fucose antibodies comigrated with the spots detected with CAH1 antibodies (see Supplemetary Information, Fig. S2c-f**),** suggesting that the mature stromal CAH1 protein is *N*-glycosylated. We further confirmed *in planta* glycosylation of CAH1 by employing the deglycosylating enzyme PNGase F. To circumvent the complication that this enzyme does not work on complex *N*-glycans harbouring α(1,3)-fucose on the proximal glucosamine residue, the experiment was carried out on a stroma fraction isolated from the *Arabidopsis mur1* mutant lacking fucose residues[15,16]. Immunoblots of both wild-type and *mur1* stroma fractions were probed with CAH1 antibodies revealing a specific CAH1 band of the expected molecular size in both cases (Fig. 3b, left panel). To establish that the *mur1* mutant provides a suitable tool for PNGase F deglycosylation experiments, and to confirm the specificity of the α(1,3)-fucose antibodies, immunodetection with α(1,3)-fucose antibodies was performed on wild-type and *mur1* mutant stroma fractions (Fig. 3b, right panel). The antibodies cross reacted with several polypeptides in the wild-type, while no signal in

the expected molecular range of CAH1 was observed in the *mur1* stroma fraction (Fig. 3b). After PNGase F treatment, the protein recognised by the anti-CAH1 antibody was completely converted into a protein exhibiting an electrophoretic mobility of the non-glycosylated polypeptide (~30 kDa) (Fig. 3c). From this we conclude that stromal CAH1 protein indeed harbours complex *N*-glycans.

β(1,2)-xylose and α(1,3)-fucose epitopes are known to be specifically added within the Golgi apparatus[13], implying that transport to and from the Golgi apparatus may represent an intermediate step in CAH1 trafficking to the chloroplast. Therefore, we analyzed the effect of brefeldin A (BFA), a fungal antibiotic that inhibits Golgi-mediated vesicular traffic[17], on the intracellular distribution of CAH1-GFP in transiently transformed *Arabidopsis* protoplasts (Fig. 4a-e). Twenty-four hours after transformation, when a fraction of the CAH1-GFP fusion protein had reached the chloroplast (Fig. 4a) BFA (180 μM) was added for 4 h. One hour prior to BFA wash-out, cycloheximide was added to block *de novo* protein biosynthesis. We observed accumulation of CAH1-GFP in ER and Golgi-like structures upon BFA treatment (Fig. 4b, c), similar to previous observations made upon BFA treatment in *Arabidopsis*[18,19]**.** Importantly, the CAH1-GFP fusion protein redistributed to the chloroplast after BFA removal. After 3 h most of the GFP fluorescence co-localized with chlorophyll autofluorescence (Fig. 4d) and after 4 h an almost complete co-localization was observed (Fig. 4e). In contrast, a GFP fusion protein targeted to the chloroplast via the Tic/Toc system was not affected by BFA treatment (see Supplementary Information, Fig. S3a, b).

We also followed the fate of endogenous CAH1 protein upon BFA treatment by employing sub-cellular fractionation studies (Fig. 4f). *Arabidopsis* suspension cells were analyzed after 4 h in the absence (lanes 1, 2, 6 and 7) and presence of BFA

(lanes 3, 4, 5, 8, 9 and 10). In the absence of BFA, the mature CAH1 form accumulated in the soluble fraction containing all the stroma content (lane 1). Under these conditions, a low molecular mass form, presumably corresponding to the unglycosylated CAH1 precursor, was found in the microsomal fraction (lane 6). Upon BFA treatment (lane 3 and 8), presence of the mature CAH1 form in the soluble fraction was reduced (lane 3) and a concomitant accumulation of partially glycosylated CAH1 forms was observed in the microsomal fraction (lane 8). Separation of this fraction from both control and BFA treated cells by sucrose density gradients showed that these CAH1 forms were localized in light dense microsomes, particularly in ER enriched fractions (data not shown). As for the CAH1-GFP localization analysis, cycloheximide was added to the cell culture after three hours of BFA incubation. After an additional hour of incubation with both drugs, BFA was washed out using a solution that still contained cycloheximide. Upon three hours after BFA wash-out, the high molecular forms of CAH1 disappeared from the microsomal fraction (lane 9), while the mature CAH1 form started to accumulate in the soluble fraction containing the stroma (lane 4). At 4h after BFA removal, the mature form of CAH1 was present in the soluble fraction and absent from the microsomal fraction (lane 5 and 10), a distribution similar to the time point before BFA treatment (lane 1 and 6). Chloroplast localized CAH1 was almost completely absent after BFA treatment, which might suggest a rapid turnover of the CAH1 protein under certain conditions. A potential regulation of CAH1 by protein degradation will thus provide an interesting objective for future studies. Taken together, these results strongly suggest that CAH1 targeting to the chloroplast involves a BFA-sensitive vesicular transport pathway.

In conclusion, our data provide strong evidence that the chloroplast proteome contains *N*-glycosylated proteins transported through the ER, in addition to the proteins synthesized in the chloroplast and those transported through the Toc/Tic apparatus. Recently, extensive studies of the chloroplast proteome have revealed the occurrence of many proteins with predicted SP inside chloroplasts[6,7]. The results presented here provide firm functional support for the existence of such a novel protein targeting pathway through the secretory system to the chloroplast in plants.

When the first genes were transferred from the genome of the cyanobacterial endosymbiont to the nuclear genome, there was probably no protein sorting system for the ancestral chloroplast. The encoded proteins may have been secreted from the eukaryotic host and subsequently taken up by the endosymbiont. Many algal groups as well as apicomplexan parasites[20,21] possess a so-called "complex" plastid that has originated from a secondary endosymbiotic event[22]. The accepted mechanism whereby proteins are routed to these "complex" plastids is through the secretory pathway[20].

As shown here, this ancestral pathway seems to have been maintained for some chloroplast proteins. The mechanism whereby these plastid proteins are translocated from the secretory pathway to the chloroplast is intriguing. In apicomplexan parasites and diatoms, plastid proteins targeted through the secretory pathway contain a bipartite N-terminal pre-sequence consisting of a SP followed by a transit peptide[23,24]. In contrast, *Arabidopsis* CAH1 is only predicted to contain an N-terminal SP, indicating that the translocation mechanism is not identical. However, the C-terminus of CAH1 is highly hydrophilic and enriched in lysine residues, like the apicomplexan transit peptides[23]. Therefore the C-terminus could possibly serve as a signal sequence in this particular uptake mechanism.

While our results strongly suggest that vesicular transport through the secretory system provides the mechanism by which CAH1 is transported into the plastid, it remains to be determined how precisely this is accomplished. For example, evidence for direct contact between the ER and the chloroplast membranes has been previously obtained[25] and biochemical interactions between the two membrane systems are essential for lipid metabolism[26]. Future detailed studies will be needed to resolve the exact mechanism used for transporting CAH1 to the chloroplast. These may further reveal whether the pathway outlined here also accommodates for chloroplast localisation of other proteins, such as rice α-amylases[8,27]. Based on indirect evidence, the authors hypothesized that vesicular transport through the Golgi apparatus is involved in the plastid targeting of these α-amylases. However, the susceptibility of the plastid isoform of α-amylase I-1 to deglycosidases[8] indicates that this protein does not traffic through the Golgi. Thus, dual import of rice α-amylases to plastids and extracellular space could occur via two mutually exclusive conserved targeting mechanisms, as it has been shown for other proteins[28]. In contrast to these previously reported findings, CAH1 is mainly localized in the chloroplast and not subjected to dual targeting. Taken together our data convincingly demonstrate transport of a protein to the chloroplast through the secretory system in plants.

METHODS

Plant material and growth conditions. Wild-type *Arabidopsis thaliana*, ecotype Columbia, and the mutant *mur1* plants were grown in a growth chamber set at 23/18°C day/night temperature, 70 % humidity and a photon flux density of 150 µmol m^{-2} s^{-1} during the 8 h photoperiod.

Cloning. A putative α-CA EST clone (*Arabidopsis thaliana*, GenBank accession number Z18493) was used to screen a total of 3.0 x 10^5 plaques from a Uni-ZAP™ XR *Arabidopsis thaliana* cDNA library (Stratagene, La Jolla, CA, USA). Nucleotide sequences of three positive clones were determined and the 5´end of the cDNA was identified through the 5´-RACE–PCR procedure (Gibco-BRL, CA, USA). A genomic library was also screened and three positive clones were subcloned. A fragment including the 5'-end of the gene and 728 bp upstream of the putative translation initiation site was sequenced.

Overexpression of recombinant CAH1 in *E. coli*. PCR was used to amplify a selected cDNA region from *CAH1* and cloned into *Bam*HI -*Xho*I digested expression vector pET23a (+) (Novagen, Madison, WI, USA). The resulting plasmid, pSLaCAH1, was verified by sequencing and encodes a recombinant *Arabidopsis* CAH1 starting at Gly(28), with an N-terminal T7-tag and a C-terminal 6-histidine tag. The construct was transformed into *E. coli* BL21 (DE3) and the expressed recombinant protein purified under denaturing conditions to near-homogeneity, using a histidine tag-binding resin, according to the pET System Manual (Novagen).

Chloroplast isolation and fractionation. Chloroplasts from the *A. thaliana* wild-type and the *mur1* mutant were purified as previously described[29]. The chloroplasts were further purified on a 50 % (v/v) Percoll gradient (Pharmacia Biotech, Sweden). Intact chloroplasts in chloroplast resuspension buffer were sonicated 3 x 30 s and centrifuged at 15,000 *g* for 30 min. The supernatant, mainly containing stroma proteins, was applied to a 1-mL MonoQ anion exchange column (HiTrap Q FF; Pharmacia) equilibrated with 20 mM Tris-HCl buffer (pH 7.8). Bound proteins were eluted with a 30-mL linear gradient from 0 to 800 mM NaCl. Each fraction was

desalted using PD-10 columns (Pharmacia). The purification process was monitored by subjecting aliquots from each fraction to western blot analyses.

Immunocytochemistry. Developing *Arabidopsis* leaves were cut into 2 mm^2 pieces in freshly-made 4% paraformaldehyde and 0.5% glutaraldehyde in a sodium phosphate sucrose buffer (0.1 M phosphate-buffer, 0.05 M sucrose, pH 7.2) and fixed for 5 h at room temperature under a gentle vacuum. After several rinses, samples were dehydrated through a graded ethanol series and embedded in LR white resin (London Resin Co Ltd, UK). Immunolocalization at the light microscope level was carried out on 1-2 mm tissue sections, cut with a diamond knife on an LKB superfrost-plus microtome and then affixed to slides. CAH1 specific antibodies at a dilution of 1:100 were used and the primary immune complexes were visualized by probing the sections for 2 h with colloidal gold-conjugated (6 nm) goat anti-rabbit IgG (diluted 1:100). The immuno-label was enhanced by using a silver enhancement kit (Biocell). Sections were then counter-stained with toluidine blue and permanently mounted for observation on a Zeiss Axiophot microscope using bright field illumination. Immunolocalization at the electron microscopy level was carried out on 150 nm ultra-thin sections picked up on uncoated 200-mesh nickel grids. The gold labelling was examined on an electron microscope after staining the grids in 2 % aqueous uranyl acetate for 10 min.

Expression in reticulocyte lysate in the presence of dog pancreas microsomes. The *CAH1* gene and the N-terminally truncated version (lacking positions 1-24) were cloned into pGEM1 (Promega, Wi ,USA) with the initiator ATG codon in the context of a "Kozak consensus" sequence[30]. The constructs were transcribed by SP6 RNA polymerase (Promega) for 1 hour at 37°C. The transcription mixture was as follows:

1-5 µg DNA template, 5 µl 10 x SP6 H-buffer (400 mM Hepes-KOH (pH 7.4), 60 mM Mg acetate, 20 mM spermidine-HCl), 5 µl BSA (1 mg/ml), 5 µl m7G (5') ppp (5') G (10mM) (Pharmacia), 5 µl DTT (50 mM), 5 µl rNTP mix (10 mM ATP, 10 mM CTP, 10 mM UTP, 5 mM GTP), 18.5 µl H_2O, 1.5 µl RNase inhibitor (50 units), 0.5 µl SP6 RNA polymerase (20 units). Translation was performed in reticulocyte lysate in the presence or absence of dog pancreas microsomes[31]. The acceptor peptide Benzoyl-NLT-methylamide (Quality Control Biochemicals Inc.) was added as a competitive inhibitor of glycosylation at a final concentration of 200 µM. For protease treatment, 5 µl 4.5 mg/ml proteinase K and 1 µl 200mM $CaCl_2$ was added to the translation mix and the sample was incubated on ice for 30 min. The treatment was stopped by addition of 0.5 ml 20 mg/ml PMSF. Translation products were analyzed by SDS-PAGE and gels were quantified on a Fuji FLA-3000 phosphoimager using Fuji Image Reader 8.1j software.

Construction of GFP reporter plasmids for transient expression in *Arabidopsis* cells. The GFP reporter plasmid, CaMV35S-sGFP(S65T), and the plasmid containing the transit peptide (TP) sequence from RBCS fused to GFP, (CaMV35S-TP-sGFP(S65T)), have been previously described[32]. DsRed2-ER protein (BD Biosciences ClonTech, Palo Alto, CA, USA) was used as an endoplasmic reticulum control. The plasmids for expression of the full or truncated *Arabidopsis* CAH1 protein fused to GFP were constructed as follows: The CaMV35S-CAH1-sGFP(S65T) corresponding to the coding region of *Arabidopsis* CAH1 was PCR-amplified using the two flanking primers for-*SalI* (TAAAAGTCGACATGAAGATTATGATGATGA) and rev1-*NcoI* (AAAACCCATGGAATTGGGTTTTTTCTTTTT) and the PCR product was cloned into the *SalI-NcoI* digested GFP reporter plasmid CaMV35S-sGFP(S65T). The protocol was similar for the other constructs. The CaMV35S-(1-40)CAH1-

sGFP(S65T) corresponding to CAH1 containing the first 40 amino acids was PCR amplified using the two flanking primers for-*SalI* and rev2-*NcoI* (GTGTCCCATGGGGTTTGGTCCATTTTTGCC). The CaMV35S-(1-40)CAH1-sGFP(S65T)-KDEL corresponding to CAH1 containing the first 40 amino acids fused to a KDEL-tagged GFP was PCR amplified using the two flanking primers for-*SalI* and rev1-*BsrGI* (TCTGCTGTACAGTCAGAGTTCATCCTTATACAGCTCGTCCATGCC). The CaMV35S-CAH1-sGFP(S65T)-KDEL corresponding to the coding region of *Arabidopsis* CAH1 fused to a KDEL-tagged GFP was PCR amplified using the two flanking primers for-*SalI* and rev1-*BsrGI*. The plasmids were sequenced to ascertain that the orientation and sequences of the inserted fragments were correct. The plasmids used for PEG transformation were prepared using the Plasmid MidiPrep kit (Bio-Rad Laboratories, CA,USA).

Transformation and confocal laser scanning microscopy. Five micrograms of the appropriate plasmid constructs were introduced into *Arabidopsis* cells using the PEG method. After transformation, cells were incubated on the plates for 24-48 h in the dark. Cells were transferred to glass slides and examined by fluorescence microscopy. Localization of GFP, GFP fusions, and DsRed2-ER was assessed in transformed cells by confocal laser scanning microscopy as outlined in the supplementary information.

ER-Tracker fluorescence analysis. *Arabidopsis* cells expressing GFP and non-transformed cells were centrifuged and resuspended in a medium containing 1 μM ER-Tracker blue-white DPX (Molecular Probes). After 45 min incubation in this medium at room temperature, cell suspensions were centrifuged and resuspended in

fresh medium without dye. Cells were transferred to glass slides and observed by confocal laser scanning microscopy as outlined in the supplementary information.

Separation of intracellular membranes by density gradient centrifugation. Isolation of the total microsome fraction and separation by density gradient centrifugation was carried out as previously described[33]. Briefly, 10 g of packed *Arabidopsis* cells frozen in liquid nitrogen were ground in a mortar and pestle, resupended in two volumes of homogenization buffer (25 mM Tris-HCl, pH 7.5, 0.25 M sucrose, 3 mM EDTA, 1 mM DTT) and centrifuged for 15 min at 10,000 *g* at 4°C. The supernatant was centrifuged for 60 min at 150,000 *g*, the supernatant (SN) collected, and the pellet (termed total microsomes) thoroughly resuspended in 1 ml of buffer containing 5 mM Tris-HCl, pH 7.5, 0.25 mM sucrose, 3 mM EDTA, and 1 mM DTT and loaded into an 11-ml linear gradient of 20% to 50% (w/w) sucrose buffered with 5 mM Tris-HCl, pH 7.5, 3 mM EDTA, and 1 mM DTT. Sucrose gradients were centrifuged at 80,000 *g* for 5 h at 4°C in a swing-out rotor (SW41 Beckman). One-milliliter fractions were collected and stored at –80°C until analysis.

Deglycosylation assays. A stroma fraction (100 µg protein/ml) enriched in CAH1 protein isolated from the mutant *mur1* of *Arabidopsis thaliana* was deglycosylated using a recombinant peptide-*N*-glycosidase F (PNGase F, Roche Diagnostics Corporation, IN,USA) according to the manufacturer instructions with some modifications. The sample was denatured at 100 °C for 5 min in the presence of 1% (w/v) SDS. After cooling the sample to room temperature, SDS was removed using the SDS-out kit (Pierce Co., Rockford, USA). The sample was then diluted with the same volume of 0.1 M Tris-HCl buffer (pH 7.8) containing 0.5% (v/v) Nonidet P-40 (Sigma, St Louis, Mo, USA). Twenty units of PNGase F were added and the sample

incubated for 24 or 48 h at 37°C. The sample was further analyzed by SDS-PAGE and immunoblotting with antibodies against CAH1. Fetuin (Sigma) was used as positive control during the deglycosylation experiments and treated identically to the stroma fraction.

Immunoprecipitation reactions. Transfected *Arabidopsis* protoplasts were resuspended in extraction buffer containing 25 mM Tris-HCl (pH 7.8), 10 mM $MgCl_2$, 5 mM EGTA, 2 mM DTT, 10% glycerol, 75 mM NaCl, 60 mM β-glycerophosphate, 0.2% Nonidet P-40, 1mM benzamidine and 1x Protease Inhibitor Cocktail (Sigma). The resuspended protoplasts were frozen in liquid nitrogen and then thawed and clarified by centrifugation. Extracts were first precleared for 2 h at 4°C with normal mouse IgG - Protein G Sepharose or normal rabbit IgG – Protein A Sepharose beads for anti-GFP (mouse monoclonal antibody sc-9996 Santa Cruz Biotechnology) and anti-CAH1 polyclonal antibodies, respectively. CAH1 fusion proteins were next precipitated by incubating the precleared extracts either with 1.5 µg of anti-GFP or anti-CAH1 antibodies for 4 h at 4°C and then capturing the immunocomplexes on 10 µl of Protein G or Protein A Sepharose beads. Beads were washed with 50 mM Tris-HCl, 150 mM NaCl, 0.1% NP 40 pH 7.5 buffer and bound proteins were eluted by boiling in 25 µl Laemmli sample buffer.

Brefeldin A treatment and recovery of transformed protoplasts and cell suspensions.

Stock solutions of brefeldin A (BFA; Sigma) were prepared at 36 mM by dissolving BFA in DMSO. Aliquots of this stock were added to 3- to 4-day-old cell suspensions and/or to 24-h-old protoplast suspensions to give a final concentration of 180 µM.

Cells were incubated with BFA for 4 h under continuous agitation. BFA-treated cells or protoplasts were harvested by low-speed centrifugation. For recovery experiments, *Arabidopsis* cells or transformed protoplasts were treated with 180 μM BFA for 4 h. Fortyfive minutes prior to the end of the incubation period cycloheximide was added to a final concentration of 50 and 100 μM for protoplasts or cells, respectively. Cells and protoplasts were washed twice with BFA-free, cycloheximide-containing culture medium, resuspended in the same medium, and samples were harvested at different times during recovery.

1. Leister, D. Chloroplast research in the genomic age. *Trends Genet.* **19,** 47-56 (2003).

2. Abdallah, F., Salamini, F. & Leister, D. A prediction of the size and evolutionary origin of the proteome of chloroplasts of *Arabidopsis*. *Trends Plant Sci.* **5,** 141-142 (2000).

3. Keegstra, K. & Cline, K. Protein import and routing systems of chloroplasts. *Plant Cell* **11,** 557-570 (1999).

4. Soll, J. Protein import into chloroplasts. *Curr. Opin. Plant Biol.* **5,** 529-535 (2002).

5. Jarvis, P. & Soll, J. Toc, Tic, and chloroplast protein import. *Biochim. Biophys. Acta* **1541,** 64-79 (2001).

6. Kleffmann, T. *et al.* The *Arabidopsis thaliana* chloroplast proteome reveals pathway abundance and novel protein functions. *Curr. Biol.* **14,** 354-362 (2004).

7. Friso, G. *et al.* In-depth analysis of the thylakoid membrane proteome of *Arabidopsis thaliana* chloroplasts: new proteins, new functions, and a plastid proteome database. *Plant Cell* **16,** 478-499 (2004).

8. Asatsuma, S. *et al*. Involvement of α-amylase I-1 in starch degradation in rice chloroplasts. *Plant Cell Physiol.* 46: 858-869 (2005).

9. Hewet-Emmett, D. & Tashian, R.E. Functional diversity, conservation, and convergence in the evolution of the α-, β-, and γ-carbonic anhydrase gene families. *Mol. Phyl. Evol.* **5,** 50-77 (1996).

10. Emanuelsson, O., Nielsen H., Brunak, S. & von Heijne, G. Predicting subcellular localization of proteins based on their N-terminal amino acid sequence. *J. Mol. Biol.* **300,** 1005-1016 (2000).

11. Nielsen, H., Engelbrecht, J., Brunak, S. & von Heijne, G. Identification of prokaryotic and eukaryotic signal peptides and prediction of their cleavage sites. *Prot. Eng.* **10,** 1-6 (1997).

12. Monné, M., Nilsson, I., Elofsson, A. & von Heijne, G. Turns in transmembrane helices: determination of the minimal length of a “helical hairpin” and derivation of a fine-grained turn propensity scale. *J. Mol. Biol*. **293,** 807-814 (1999).

13. Lerouge, P. *et al.* N-glycosylation biosynthesis in plants: recent developments and future trends. *Plant Mol. Biol.* **38,** 31-48 (1998).

14. Faye, L., Gomord, V., Fitchette-Laine, A.C. & Chrispeels, M.J. Affinity purification of antibodies specific for Asn-linked glycans containing α 1-3 fucose or β 1-2 xylose. *Anal. Biochem.* **209,** 104-108 (1993).

15. Rayon, C. *et al.* Characterization of *N*-glycans from *Arabidopsis thaliana*. Application to a fucose-deficient mutant. *Plant Physiol.* **119,** 725-734 (1999).

16. Bonin, C., Potter, I., Vanzin, G.F. & Reiter, W.D. The *MUR1* gene of *Arabidopsis thaliana* encodes an isoform of GDP-D-mannose-4,6-dehydratase, catalyzing the first step in the *de novo* synthesis of GDP-L-fucose. *Proc. Natl. Acad. Sci. USA* **94,** 2085-2090 (1997).

17. Ritzenthaler, C. *et al.* Reevaluation of the effects of brefeldin A on plant cells using tobacco bright yellow 2 cells expressing Golgi-targeted green fluorescent protein and COPI antisera. *Plant Cell* **14,** 237-261 (2002).

18. Lee, M.H. *et al.* ADP-ribosylation factor 1 of Arabidopsis plays a critical role in intracellular trafficking and maintenance of endoplasmic reticulum morphology in Arabidopsis. *Plant Physiol.* **129,** 1507-1520 (2002).

19. Zheng, H., Kunst, L., Hawes, C. & Moore, I. A GFP-based assay reveals a role for RHD3 in transport between the endoplasmic reticulum and Golgi apparatus. *Plant J.* **37,** 398-414 (2004).

20. Sulli, C. & Schwartzbach, S.D. The polyprotein precursor to the *Euglena* light-harvesting chlorophyll a/b-binding protein is transported to the Golgi apparatus prior to chloroplast import and polyprotein processing. *J. Biol. Chem.* **270,** 13084-13090 (1995).

21. Waller, R.F., Reed, M.B., Cowman, A.F. & McFaden, I. Protein trafficking to the plastid of *Plasmodium falciparum* is via the secretory pathway. *EMBO J.* **19,** 1794-1802 (2000).

22. Delwiche, C. Tracing the thread of plastid diversity through the tapestry of life. *Am. Nat.* **154,** S164-S177 (1999).

23. Foth, B.J. *et al.* Dissecting apicoplast targeting in the malaria parasite *Plasmodium falciparum*. *Science* **299,** 705-708 (2003).

24. Kilian, O. & Kroth, P.G. Identification and characterization of a new conserved motif within the presequence of proteins targeted into complex diatom plastids. *Plant J.* **41,** 175-183 (2005).

25. Moreau, P. *et al.* Lipid trafficking in plant cells. *Prog .Lipid Res.* **37,** 371-391 (1998).

26. Xu, C., Fan, J., Riekhof, W., Froehlich, J.E. & Benning, C. A permease-like protein involved in ER to thylakoid lipid transfer in *Arabidopsis*. *EMBO J.* **22,** 2370-2379 (2003).

27. Chen, M-H., Huang, L-F., Li, H-m., Chen, Y-R. & Yu, S-M. Signal peptide-dependent targeting of a rice α-amylase and cargo proteins to plastids and extracellular compartments of plant cells. *Plant Physiol.* **135,** 1367-1377 (2004).

28. Levitan, A. *et al.* Dual targeting of the protein disulfide isomerase RB60 to the chloroplast and the endoplasmic reticulum. *Proc. Natl. Acad. Sci. USA* **102,** 6225-6230 (2005).

29. Kunst, L. In: Methods in Molecular Biology, Vol 82. *Arabidopsis* protocols, Martinez-Zapater, J. & Salinas, J., eds (Humana Press Inc., Totowa, New Jersey, USA, 1998), pp 43-53.

30. Kozak, M. Regulation of translation in eukaryotic systems. *Annu. Rev. Cell. Biol.* **8,** 197-225 (1992).

31. Hermansson, M., Monné, M. & von Heijne, G. Formation of helical hairpins during membrane protein integration into the endoplasmic reticulum membrane. Role of the N- and C-terminal flanking regions. *J. Mol. Biol.* **313,** 1171- 1179 (2001).

32. Chiu, W-L. *et al.* Engineered GFP as a vital reporter in plants. *Curr. Biol.* **6,** 325-330 (1996).

33. Gasparian, M. *et al.* Identification and characterization of an 18-kilodalton, VAMP-like protein in suspension-cultured carrot cells. *Plant Physiol.* **122,** 25-33 (2000).

Acknowledgements

We thank Alexandra Kraut for technical assistance in the transient expression of GFP fusions in plant cells; Loïc Faye for the gift of antibodies against xylose and fucose residues; Colin Robinson for helpful discussion; and Judy Brangeon and Roland Boyer for technical help in the immunocytochemistry experiments. The authors are grateful to B Martin for critical reading of the manuscript. This work was supported by grants from the Swedish National Research Council, FORMAS, Wallenberg and Kempe Foundations, and the Swedish Foundation for Strategic Research.

Correspondence and requests for materials should be addressed to G.S. (e-mail: goran.samuelsson@plantphys.umu.se).

Note: supplementary information is available at the Nature Cell Biology website.

LEGENDS

Fig. 1. The deduced amino acid sequence of CAH1 (**a**). Red, blue and green coloured amino acid residues corresponds to 22 out of 36 putative active site

amino acids in α-type carbonic anhydrases. Blue histidine residues denote Zinc binding ligands, conserved in all active α-CAs. Green residues are conserved in the active site of α-type CAs while red residues are conserved but found outside the active site. The predicted signal peptide is shown in italics and the predicted processing site indicated by an arrow. Underlined triplets indicate possible *N*-glycosylation sites. CAH1 is localized in the *Arabidopsis* chloroplast stroma. (**b** and **c**) Immunogold labeling of CAH1 in *Arabidopsis* leaves, indicating that this protein is located in the chloroplast stroma. Sections were examined under the light microscope (**b**), scale bar 20 μm; or electron microscope (**c**), the arrows indicate gold particles, scale bar 100 nm. (**d**) Coomassie staining (top panel) and immunoblot with CAH1 antibodies of total leaf extracts and different chloroplast sub-fractions: CE, total leaf extract; Cp, total chloroplast extract; E, chloroplast envelope membranes; S, stroma; T, thylakoid membranes. (**e, f** and **g**) Transient expression of the GFP-tagged form of the *Arabidopsis* CAH1 protein in *Arabidopsis* cells under the 35S promoter, scale bar 10 μm (**e**). As subcellular localization controls, we used ER-Tracker (Molecular Probes) staining, the 35S-sGFP(S65T) plasmid (GFP), scale bar 10 μm (**f**) and the 35S-TP-sGFP(S65T) plasmid (TP-GFP) containing the transit peptide (TP) sequence from RbcS fused to GFP, scale bar 10 μm (**g**).

Fig. 2. CAH1 is taken up into the ER and glycosylated. (**a** and **b**) Retention in the ER of KDEL-tagged CAH1-GFP fusion proteins, transiently expressed in *Arabidopsis* cells, scale bar 10 μm. As subcellular localization controls, cells were stained with a dye selective for the ER (ER-Tracker, Molecular Probes).

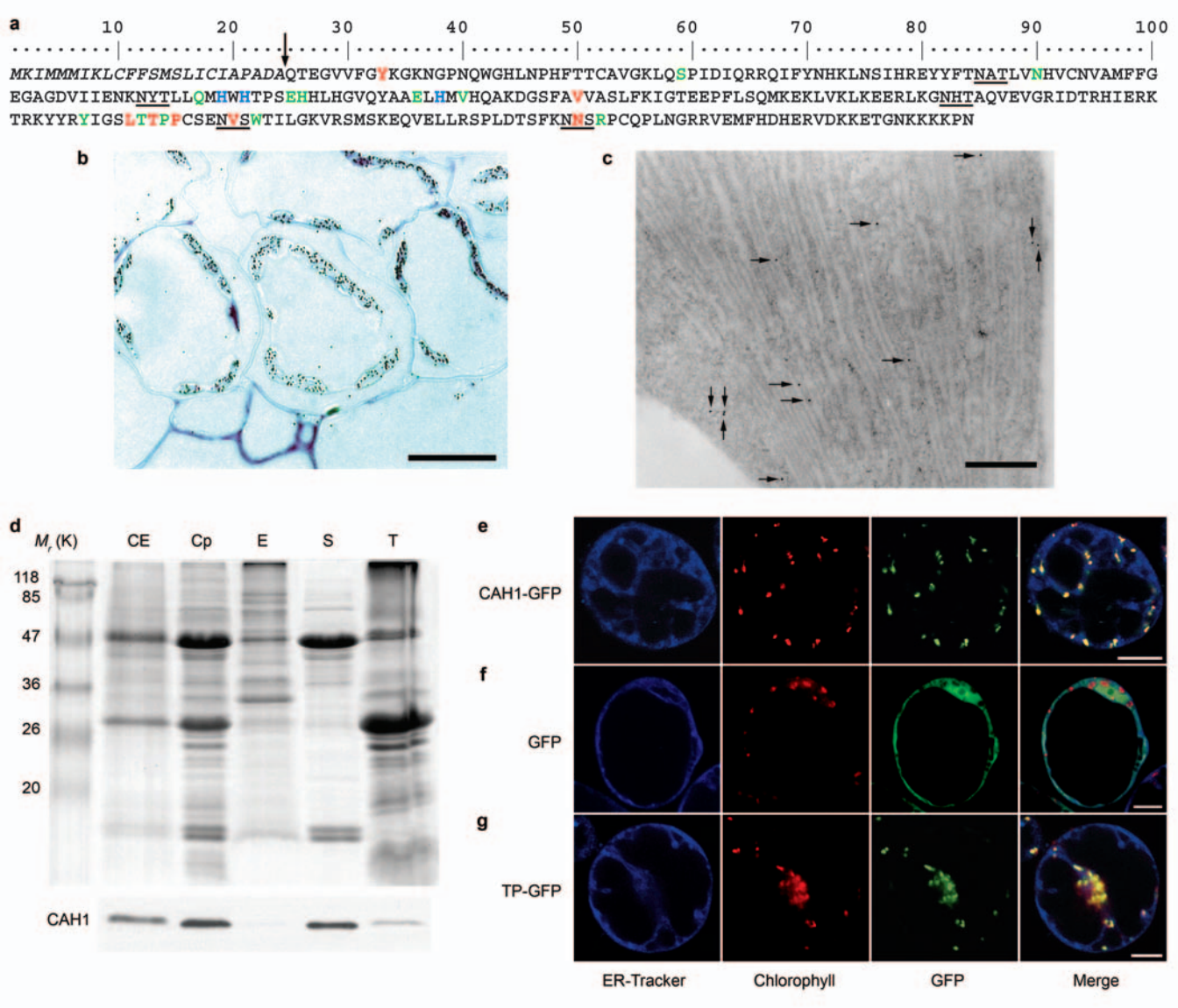
(**a**) The first 40 amino acid residues of CAH1, containing the signal peptide (SP) for the ER, were fused to the N-terminus of a KDEL-tagged GFP. (**b**) The full-length CAH1 was fused to the N-terminus of a KDEL-tagged GFP. (**c**) *In vitro* uptake into dog pancreas microsomes. CAH1 (lanes 1, 2, 3, 5, 6, and 7) and a truncated version lacking the signal peptide (SP) (lanes 4, 8, 9, and 10) were expressed *in vitro* in the absence (-) and presence (+) of rough dog pancreas microsomes (RM) either with no additions (lanes 1, 2, 4, 5, 6, 8, and 9), with addition of the competitive glycosylation inhibitor benzoyl-Asn-Leu-Thr-methylamide (AP) (lane 3), or with addition of protease (PK) (lanes 7 and 10). The precursor and the processed form of CAH1 are indicated by white and black dots, respectively. G, *N*-glycosylated isoforms. The wild-type CAH1 is taken up into the microsomes, glycosylated, and the signal peptide cleaved off. The truncated form of CAH1, lacking the signal peptide, is not taken up into the microsomes.

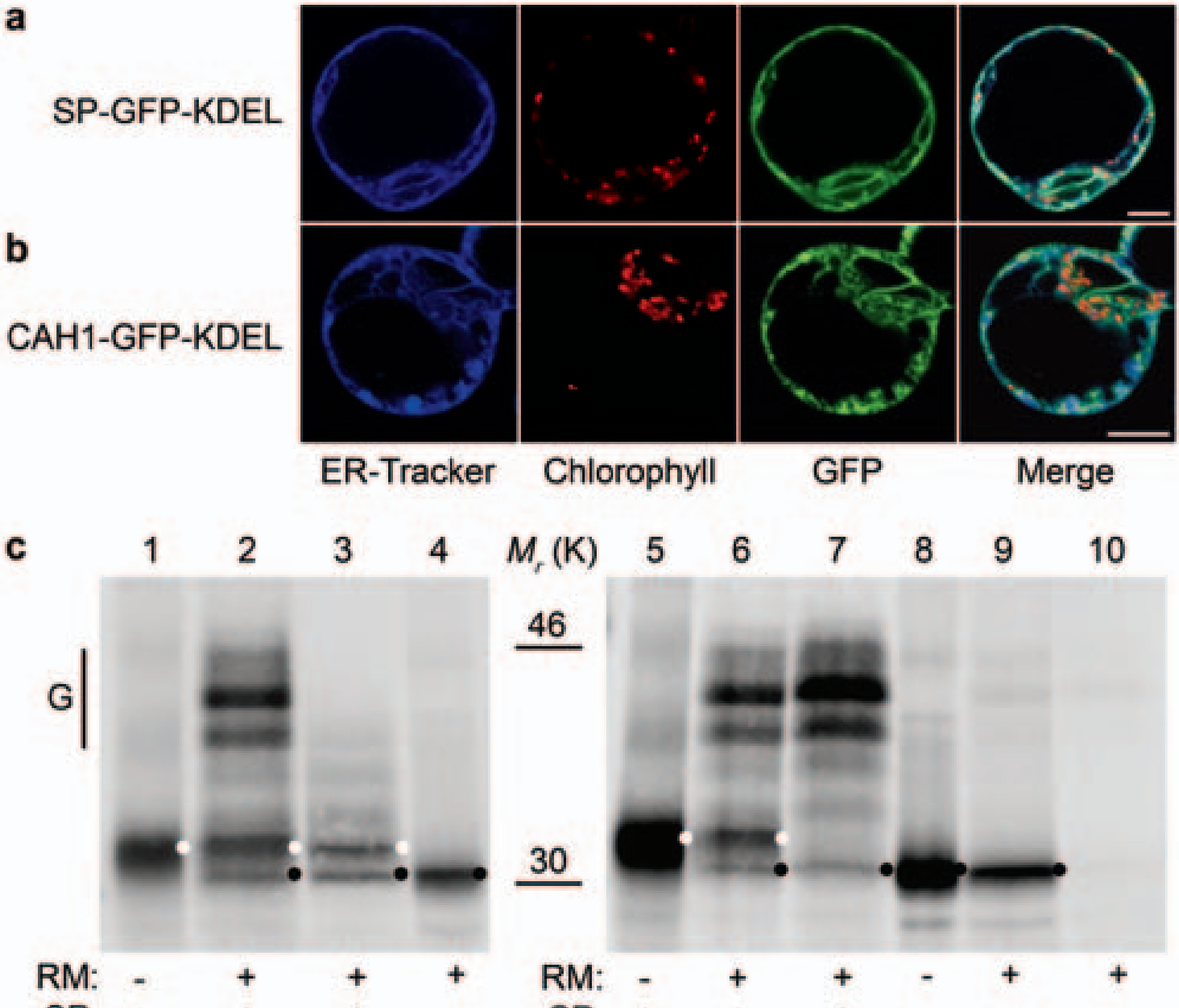
Fig. 3. Chloroplast stroma contains an *N*-glycosylated isoform of CAH1. (**a**) Immunoblot with antibodies against α(1-3) fucose residues of CAH1-GFP fusion protein immunoprecipitated with GFP antibodies from extracts of transformed *Arabidopsis* protoplasts. (**b**) Stroma fractions from wild-type and mutant *mur1* of *Arabidopsis* were separated by SDS-PAGE electrophoresis and immunoblotted with antibodies against CAH1 (CAH1) and α(1,3)-fucose epitopes (Fucose) at a dilution of 1:1000. (**c**) Deglycosylation of the CAH1 fraction leads to a shift in the electrophoretic mobility of CAH1 protein. A stroma fraction enriched in CAH1 polypeptide was deglycosylated for 24 or 48

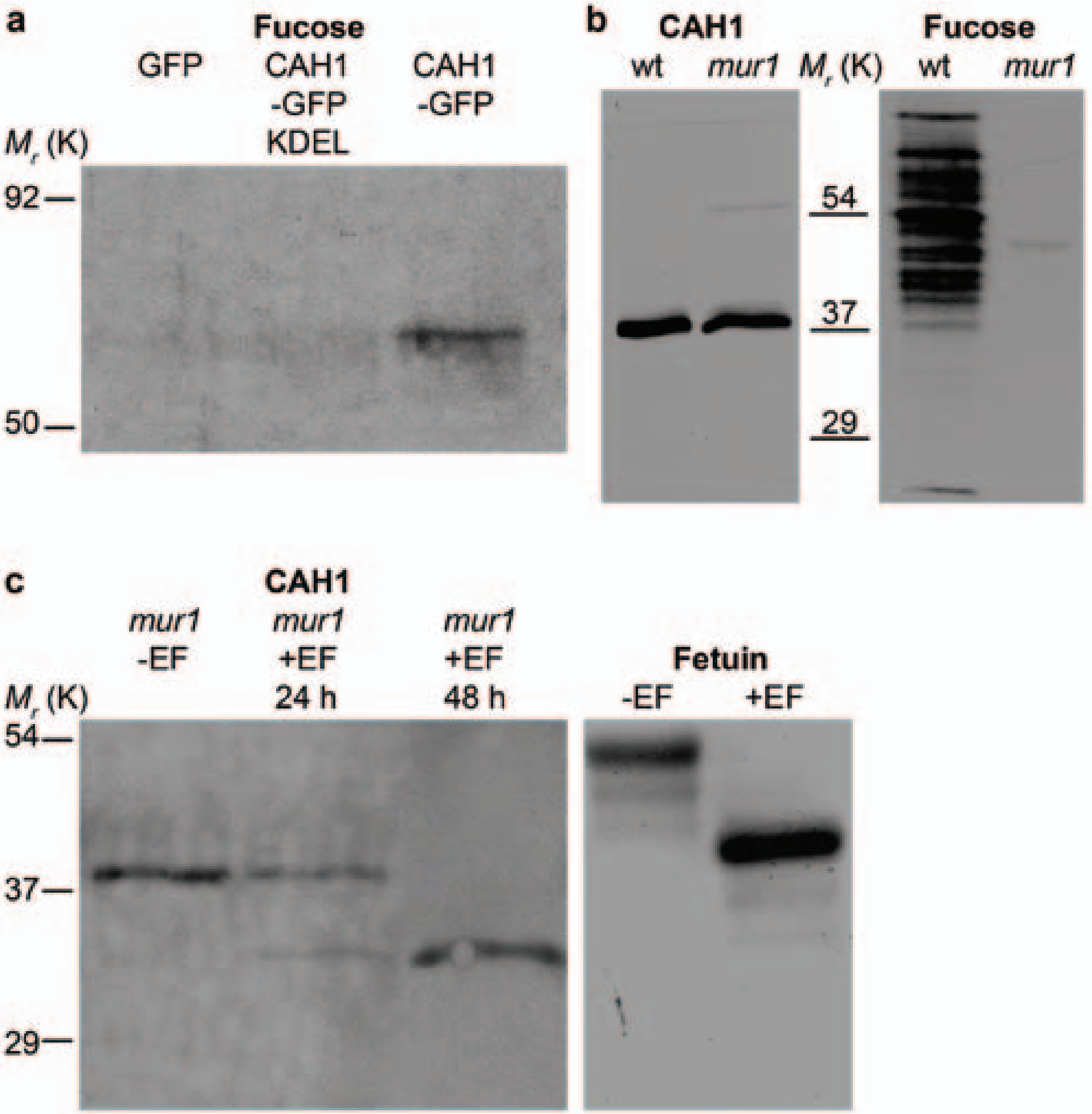
h using PNGase F (EF). Fetuin (Sigma) was used as a control for the deglycosylation reaction.

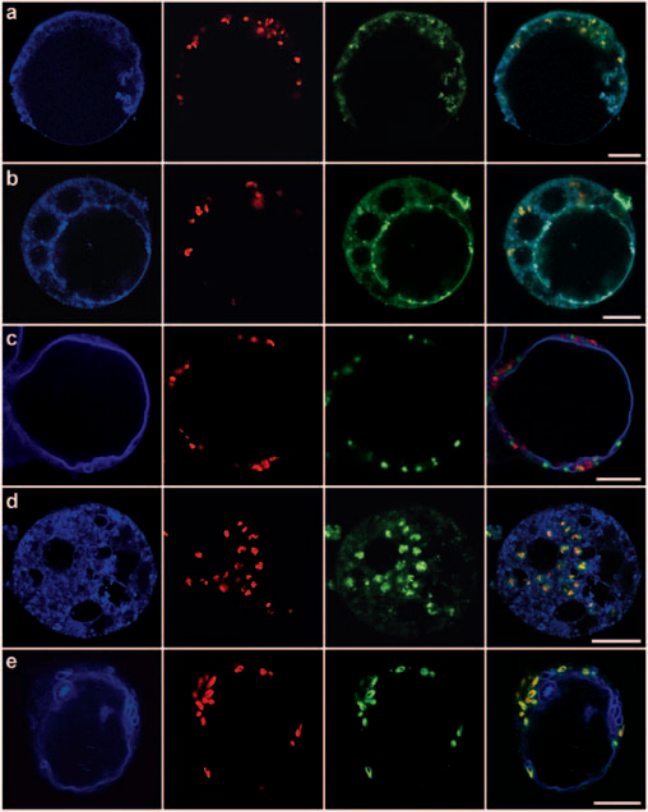
Fig. 4. Effect of BFA on chloroplast targeted CAH1-GFP fusion construct in *Arabidopsis* protoplasts (**a-e**) and native CAH1 in *Arabidopsis* cell suspensions (**f**). (**a**) Twenty four hours after transformation, CAH1-GFP is mainly localized in the ER but also begins to reach the chloroplast, as revealed by colocalization of the fluorescence with that of the ER-Tracker and chlorophyll. (**b** and **c**) Treatment with 180 μM BFA for 4 h causes the redistribution of CAH1-GFP to ER and Golgi-like structures. (**d**) Three hours after removal of BFA and in the presence of 50 μM cycloheximide, CAH1-GFP fusion was mainly relocated to the chloroplast. (**e**) Four hours after removal of BFA and in the presence of 50 μM cycloheximide, CAH1-GFP fusion was completely located in the chloroplasts, scale bar 10 μm. (**f**) Effect of BFA on the distribution of native CAH1 between soluble, stroma containing (lanes 1-5) and microsome enriched fractions (lanes 6-10) of *Arabidopsis* suspension cells. *Arabidopsis* cell suspensions were treated for 4 h in the absence (lanes 1, 2, 6, and 7) and presence (lanes 3, 4, 5, 8, 9, 10) of 180 μM BFA. Control cells were incubated for 5 h without (lanes 1 and 6) and with (lanes 2 and 7) 100 μM cycloheximide. After 4 h of treatment with BFA CAH1 disappeared from the supernatant, while different isoforms accumulated in the microsome fraction (lanes 3 and 8). At this time BFA was removed and cells were grown in cycloheximide contaning media (added 1 h before removal of BFA) for 3 h (lanes 4 and 9) and 4 h (lanes 5 and 10). The mature CAH1 form is completely relocated in the stroma fraction 4 h after removal of BFA. All the

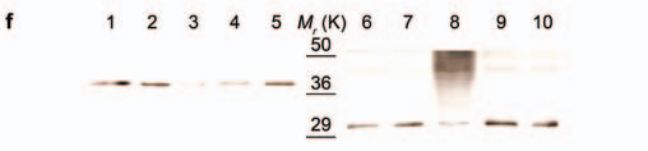
samples were immunoblotted with antibodies against CAH1. Antimycine A resistant NADH cytochrome *c* reductase activity, a marker for ER, was measured in both fractions (see Supplementary Information, Fig. S3c).


a
10 20 30 40 50 60 70 80 90 100
MKIMMMIKLCFFSMSLICIAPADAQTEGVVFGYKGKNGPNQWGHLNPHFTTCAVGKLQSPIDIQRRQIFYNHKLNSIHREYYFTNATLVNHVCNVAMFFG
EGAGDVIIENKNYTLLQMHWHTPSEHHLHGVQYAAELHMVHQAKDGSFAVVASLFKIGTEEPFLSQMKEKLVKLKEERLKGNHTAQVEVGRIDTRHIERK
TRKYYRYIGSLTTPPCSENVSWTILGKVRSMSKEQVELLRSPLDTSFKNMSRPCQPLNGRRVEMFHDHERVDKKETGNKKKKPN
b
c
d
Mr (K)
CE
Cp
E
S
T
118
85
47
36
26
20
CAH1
e
CAH1-GFP
f
GFP
g
TP-GFP
ER-Tracker
Chlorophyll
GFP
Merge


a
SP-GFP-KDEL
b
CAH1-GFP-KDEL
ER-Tracker
Chlorophyll
GFP
Merge
c
1
2
3
4
M_r (K)
5
6
7
8
9
10
46
30
G
RM: - + + + RM: - + + - + +
SP: + + + - SP: + + + - - -
AP: - - + - PK: - - + - - +


a
Fucose
GFP
CAH1 -GFP KDEL
CAH1 -GFP
M_r (K)
92
50
b
CAH1
wt
mur1
M_r (K)
54
37
29
Fucose
wt
mur1
c
CAH1
mur1 -EF
mur1 +EF 24 h
mur1 +EF 48 h
M_r (K)
54
37
29
Fetuin
-EF
+EF


a
b
c
d
e
ER-Tracker
Chlorophyll
GFP
Merge


f
1
2
3
4
5
M_r (K)
6
7
8
9
10
50
36
29