Cytidine Deaminase Motifs within the DYW Domain of Two Pentatricopeptide Repeat-containing Proteins Are Required for Site-specific Chloroplast RNA Editing*

Received for publication, October 27, 2014, and in revised form, December 12, 2014 Published, JBC Papers in Press, December 15, 2014, DOI 10.1074/jbc.M114.622084

Jessica A. Wagoner, Tao Sun, Lin Lin, and Maureen R. Hanson

From the Department of Molecular Biology and Genetics, Cornell University, Ithaca, New York 14853

Background: Pentatricopeptide repeat (PPR) proteins are site-specific RNA editing factors; many have C-terminal motifs of unknown function.

Results: Site-directed mutagenesis of two PPR-DYW domains significantly reduces editing efficiency.

Conclusion: Residues conserved in cytidine deaminase in the DYW domains are important for editing activity.

Significance: This study strengthens the evidence that the DYW domains of PPR proteins carry the deaminase activity necessary for C-to-U modification.

In angiosperm organelles, cytidines are converted to uridines by a deamination reaction in the process termed RNA editing. The C targets of editing are recognized by members of the pentatricopeptide repeat (PPR) protein family. Although other members of the editosome have begun to be identified, the enzyme that catalyzes the C-U conversion is still unknown. The DYW motif at the C terminus of many PPR editing factors contains residues conserved with known cytidine deaminase active sites; however, some PPR editing factors lack a DYW motif. Furthermore, in many PPR-DYW editing factors, the truncation of the DYW motif does not affect editing efficiency, so the role of the DYW motif in RNA editing is unclear. Here, a chloroplast PPR-DYW editing factor, quintuple editing factor 1 (QED1), was shown to affect five different plastid editing sites, the greatest number of chloroplast C targets known to be affected by a single PPR protein. Loss of editing at the five sites resulted in stunted growth and accumulation of apparent photodamage. Adding a C-terminal protein tag to QED1 was found to severely inhibit editing function. QED1 and RARE1, another plastid PPR-DYW editing factor, were discovered to require their DYW motifs for efficient editing. To identify specific residues critical for editing, conserved deaminase residues in each PPR protein were mutagenized. The mutant PPR proteins, when expressed in qed1 or rare1 mutant protoplasts, could not complement the editing defect. Therefore, the DYW motif, and specifically, the deaminase residues, of QED1 and RARE1 are required for editing efficiency.

Vascular plants modify cytidines post-transcriptionally to uridines in chloroplasts and mitochondria in a process termed RNA editing. Cytidine can be converted to uridine by deamination; however, the identity of the cytidine deaminase that carries out the reaction has remained elusive. The plant organelle editosome is a small protein complex of 200–400 kDa (1) whose entire composition is not yet known, although members of several Arabidopsis protein families have been found to be essential for efficient RNA editing (1–6). One hypothesis is that deaminase activity is provided by the DYW domain that is present C-terminally on many, but not all, pentatricopeptide (PPR)2 motifs-containing editing factors. The DYW domain contains the HXE and CXXCH motifs that are conserved in cytidine deaminases (7). Short cis-elements upstream of C targets of editing are known to bind PPR proteins, which evidently serve as the recognition factors that determine which C will undergo editing (8–10). Recently, a recognition code for the interaction of PPR motifs with particular RNA nucleotides has been proposed (11–13).

Many PPR editing factors recognize only one editing site, but there are a number that can recognize multiple editing sites. For example, CLB19 is involved in the editing of two C targets in the chloroplast, rpoA C200 and clpP C559, whereas CRR22 and OTP84 can each recognize three different C targets in the chloroplast (14, 15). We describe here a PPR-DYW editing factor, quintuple editing 1 (QED1), that affects five different chloroplast C targets, currently the largest number of chloroplast sites affected by a known PPR protein. We observed that the DYW domains of QED1 and of RARE1, another PPR-DYW protein we have previously studied (16), are essential for editing of their respective C targets. To discern the possible role of the cytidine deaminase signatures HXE and CXXCH found in the DYW domain of PPR proteins, we performed site-directed mutagenesis and assayed the effect on editing in transfected mutant protoplasts. We also investigated whether the sequences comprising a putative second zinc-binding site in the DYW domain (17) were necessary for RNA editing.

* This work was supported by National Science Foundation (NSF) Grants MCB-0929423 and MCB-1330294 (to M. R. H.).

1 To whom correspondence should be addressed: Dept. of Molecular Biology and Genetics, Cornell University, Biotechnology Bldg., Ithaca, NY 14853. Tel.: 607-254-4833; Fax: 607-255-6249; E-mail: m rh5@cornell.edu.

2 The abbreviations used are: PPR, pentatricopeptide repeat; AD, activation domain; BD, binding domain; QED1, quintuple editing factor 1; PPE, poisoned primer extension; RARE1, required for accD RNA editing; RIP, RNA editing factor-interacting protein; RRM, RNA recognition motif; nt, nucleotides; VIGS, virus-induced gene silencing.
EXPERIMENTAL PROCEDURES

Arabidopsis Lines—The T-DNA insertion line CSHL_GT13864 (Landsberg erecta ecotype) was obtained from the Arabidopsis Biological Resource Center (ABRC), and the presence of the T-DNA insert was verified through PCR with primers 5′/H11032-GCTTGCATAGTTGATGTTCTG-3′ and the Ds3-1 primer 5′/H11032-ACCCGACCGGATCGTATCGGT-3′ (18) using Bio-Rad Red master mix. The T-DNA insertion line SALK_092402C (Columbia-0 ecotype) was also obtained from the ABRC, and homozygosity of the mutants was confirmed. Arabidopsis seeds were imbibed in water at 4 °C for 3 days and then seeded on MetroMix. Plants were grown in both long day (16-h light, 8-h dark, 79 μmol/m²/s of light, 22°C) and short day (10-h light, 14-h dark, 116 μmol/m²/s of light, 22°C) conditions.

Complementation and Mutational Analysis—RARE1 and the 2217-bp wild-type coding region of QED1 were amplified using Phusion polymerase and TOPO-cloned into PCR8/GW/TOPO (Invitrogen). These constructs were used in LR Clonase II recombination reactions with pEXSG-EYFP (19) and a modified pBI121 vector (1, 20) to generate the full-length, DYW, E, and E truncated constructs driven by a 35S promoter. Site-directed mutagenesis of conserved deaminase residues was performed by two-step PCR. Both the full-length construct and the mutagenized product were digested with BamHI and SacI, and the mutagenized product was ligated into the digested full-length construct. Oligonucleotides used for mutagenesis are listed in Table 1.

Generation of Transgenic Plants—A construct in the modified pBI121 vector carrying the QED1 wild-type coding region

| Primer Name | Sequence 5′-3′ | Purpose |
|-------------|----------------|---------|
| QED1 start  | ATGGCTATCTTCTCCACACGAC | amplification |
| QED1 stop   | GAAATATCCAGAAATCGTTGTTAC | amplification |
| QED1 5′2531 | GCTTGCAATCGTATGCTTCCGCT | genotyping |
| Ds3.1       | ACCGGACCGGATCGTATCGGT | genotyping |
| QED1 DWstop | CTACAGTCTTCATACACCTGTTG | amplification |
| QED1 Eplusstop | CTACTCATGATCTGACCATCGA | amplification |
| QED1 Edomainstop | CTAAGTCTGGAGAATCCGTCA | amplification |
| QED1 BamHI  | TAGTTAGATGAGATCCATAGCTAC | cloning |
| pexyfp saci R | GGACTCTAGAGCTTCTACTGT | cloning |
| QED1 SalI-mat | TATATAGTCGACCATGTCATCCGCTAC | cloning |
| QED1 stop-Spel | GACTTCCACTAGTCCAGAAATCGTTACGG | cloning |
| QED1 CxxChmutF | ATTAAGAAATCTAAGGGTGGTCAGCTGCTGCTGTCTCGATGCTAA | mutagenesis |
| QED1 CxxChmutR | TTACCAGCTGAACGAGCTGCAGCCACCTTAGATCTTCTAAT | mutagenesis |
| QED1 HSEmutF | GAACACTCTCATAATTCTCGACGAGCAAAGGTGGCTATTTGTTTAT | mutagenesis |
| QED1 HSEmutR | ATAACAAATAGCCAGCTTTGCTGCTGAGATTTGAGAATTTGTTTAC | mutagenesis |
| qed1 CC sermutF | ATTAAGAATCTAGAGGGTGCTCTGAGATTCTCAGTCTCCGCTAA | mutagenesis |
| qed1 CC2 sermutR | TTAGCCACTGAAATGAGATCTCAGACACACCATGATTGTTTAA | mutagenesis |
| QED1 PGC F  | GAGTCACTGACAGTTAAGAAAGAAGAGCCAGAGCCACGCTCAAATGAA | mutagenesis |
| QED1 PGC R  | CATCGATTTCAATTAGCTGCTGCTTCTCTTCATTAAGTCCAGT | mutagenesis |
| QED1 CCend mutF | CATCATTTCAAGAAATGGCGAGCTTCGCTAAGATTTCTGTTGTA | mutagenesis |
| QED1 CCend mutR | GAAATATCCAGAAATCGTTAGGGAAGCTGCTCACATGATTGTTTCA | mutagenesis |
| RARE1 start | CTATGACGATTCTCAGTGAC | amplification |
| RARE1 fullstop | CATCATCAGTCACCAGTATCG | amplification |
| RARE1 dywstop | CTACATGAAACCACATCAACTCTCT | amplification |
| RARE1 5′1566 | GCAATATCAGCTCCGGAATGCTCA | sequencing |
| RARE1 HSEmutF | GAGAAGACGCTTCTTATGCTGCTGCAGACCTGCATTGCTAC | mutagenesis |
| RARE1 HSEmutR | ATGCAGATGCAAGCTCAGCATCAAGAAGCTGTTTCTC | mutagenesis |
| RARE1 CCsermutF | TTTCAAGAAATCTCGGGCAAGCGCCGAGATGATGCTGAGTTGAAAGCAT | mutagenesis |
| RARE1 CCsermutR | ATGCAGTTGCAAACTCATGACATCTACTGGCTGGCTCGCGAAGATTTT | mutagenesis |
was transformed into *Agrobacterium tumefaciens* GV3101. Floral dip transformation of *ged1* CSHL_GT13864 homozygous mutant plants was performed as in Ref. 21.

**Protoplast Transfection**—The cellulase/maceroyzyme solution was prepared following the protocol of Yoo *et al.* (22) and contained 1.5% cellulase R10, 0.4% macerozyme R10, 0.4 mM mannitol, 20 mM KCl, 20 mM MES, pH 5.7, 10 mM CaCl₂, 0.1% BSA. *Arabidopsis* plants were grown in short days (8 h of light). Fully expanded leaves were sliced into 0.5–1-mm strips using razor blades with minimal wounding and immersed in the enzyme solution. The digestion mixture was vacuum-infiltrated with minimal wounding and immersed in the enzyme solution. The digestion mixture was vacuum-infiltrated for 2 min and incubated overnight in the dark without shaking. The protoplasts were harvested in round-bottomed tubes by 2 min of centrifugation at 300 × g. The supernatant was discarded, and the cells were resuspended in the W5 solution (154 mM NaCl, 125 mM CaCl₂, 5 mM KCl, and 2 mM MES, pH 5.7). The protoplasts were left on ice for 30 min before 2 min of centrifugation at 300 × g. The supernatant was discarded, and MMG solution (0.4 mM mannitol, 15 mM MgCl₂, 4 mM MES, pH 5.7) was used to resuspend the protoplasts. Cell number was counted using a hemocytometer. The final protoplast density was adjusted to 2 × 10⁵/ml. 30 μg of plasmid DNA was used to transfect 300 μl (6 × 10⁴) protoplasts as in Jen Sheen’s protocol (23). After a 30-min incubation, a 3× volume of W5 solution was added into each tube to stop the transfection. The cells were spun down by 2 min of centrifugation at 300 × g and then resuspended in 1 ml of W5 solution. Transfected protoplasts were cultured for 3 days in 24-well tissue culture plate in the dark at room temperature to allow accumulation of protein and edited transcripts before RNA extraction. A gene encoding GFP targeted to the endoplasmic reticulum (24) was also transfected in a sample to monitor transfection efficiency.

**Measuring Editing Extent**—RNA extraction and RT-PCR was performed as described previously (25). Poisoned primer extension (PPE) editing assays of virus-induced gene silencing (VIGS) and QED1 mutant lines was carried out as described by Peeters and Hanson (26), except with the use of fluorescently labeled primer (16). Assays of editing efficiency in protoplasts were performed with the amplification primers listed in Robins *et al.* (16).

**Yeast Two-hybrid Assays**—Mature coding sequences of QED1, RIP1, RIP2, RIP3, RIP5, and RIP9 were amplified using Phusion polymerase and cloned into the PCR8/TOPO/GW vector. These constructs were used in LR Clonase II recombination reactions with pGADT7/GW and pGBKTK7/GW vectors to fuse the GAL4 activation domain (AD) and DNA binding domain (BD), respectively, to the N termini of the mature coding sequences. The yeast two-hybrid assay was performed as described in Ref. 3.

**Protein Extraction and Immunoblotting**—Samples were stored at −20 °C before polyacrylamide gel electrophoresis. 6× SDS loading buffer (300 mM Tris·HCl, pH 6.8, 60 mM EDTA, 12% SDS, 36% glycerol, 6% β-mercaptoethanol, 0.024% bromophenol blue) was added to samples and heated before loading onto a precast gradient SDS-PAGE gel (Bio-Rad Mini-PROTEAN TGX Any kD). After transfer to a nitrocellulose membrane, protein was probed with anti-FLAG antibody (Sigma-Aldrich).

**RESULTS**

**Identification of a Candidate Editing Factor**—We identified the PPR-encoding gene At2g29760 as a candidate editing factor by a previously described candidate gene approach (16). Briefly, the *Arabidopsis* genome was surveyed to identify genes encoding a PPR protein carrying a DYW motif, as a number of known editing factors are PPR-DYW proteins. These genes were then screened to identify ones that were predicted to be chloroplast-targeted according to both Target P and Predator. The predicted chloroplast-targeted proteins were used in reciprocal best-hit analysis with the *Arabidopsis* and rice nuclear genome to identify ones that had no rice orthologs. Such genes were therefore candidates for editing of C targets that occur in *Arabidopsis* but not rice. RARE1, which encodes a PPR-DYW editing factor for a site in the *Arabidopsis* accD transcript, and which does not exist in rice, was previously identified by this strategy (16).

**Expression of AT2G29760 Is Required for the Editing of Five Sites on Five Different Chloroplast-encoded Transcripts**—To determine whether At2g29760 is required for C-to-U editing events at one or more of the known chloroplast editing sites, VIGS trials were carried out. Silencing of At2g29760 was achieved by introducing amplified gene-specific DNA (Fig. 1), as defined in the Complete *Arabidopsis* Transcriptome MicroArray (CATMA) database, in a modified co-silencing strategy (16). The vector allows direct visual identification of tissue exhibiting GFP silencing. In GFP-silenced plants, 5 of 34 sites assayed by PPE, on five different transcripts, exhibited significantly lower levels of editing: accD C1568, matK C640, ndhB C872, rpoB C2432, and rps12 i1C58. Notably, all five of these sites are either absent or not edited in rice. Editing efficiency of *accD* C794, in *cis* of *accD* C1568, was not impacted by the silencing of At2g29760, nor were sites in *cis* relative to *ndhB* C872 and *rpoB* C2432 (data not shown). Because five editing sites are affected, the gene was named QED1 for quintuple edit-
qed1 is a PPR protein containing 14 PPR motifs organized into four PLS blocks as well as an E/E+ and DYW motif (Fig. 1).

Homozygous mutant lines SALK_092402C and CSH-L_GT13864 were acquired to assay C-to-U editing at each of the five sites identified by VIGS. Plants of line SALK_092402C (ecotype Columbia) were genotyped and determined to have a T-DNA insertion positioned 15 nucleotides upstream of the QED1 start codon (Fig. 1). The 5′-UTR is at least 164 nt, as a database search located a QED1 expressed sequence tag that was created by 5′ rapid amplification of cDNA ends (expressed sequence tag EG446305.1) that extends 164 nt upstream of the QED1 start codon. The CSHL line contains an insertion in the E domain (Fig. 1). No QED1 RNA was detected in the 2-week old seedling rosette leaves of the CSHL line after 40 cycles of RT-PCR (data not shown). Homozygous mutant plants carrying either mutant allele demonstrated a significant reduction in editing efficiencies of the same five sites found by gene silencing (Fig. 2). No editing was detected at any of the five sites in the CSHL plants, which carry the coding region insertion. As this mutant allele exhibits a stronger editing phenotype than the other allele, subsequent experiments utilized the CSHL line. To confirm that the qed1 allele in the CSHL line specifies the defects in RNA editing, we introduced the wild-type coding region into the mutant background. Editing of all five sites was restored, as can be seen in PPE assays of two representative plant transformants (Fig. 3).

Editing of three (matK C640: H-to-Y; ndhB C872: S-to-L; rpoB C2432: S-to-L) of the five sites leads to amino acid changes.
The QED1 sequences and cis-element sequences were also analyzed with respect to the PPR code proposed by Yagi et al. (13) where, in addition to the amino acids at the 6 and 1' position (referred to as 4 and ii in Ref. 13), a third amino acid at position 3 (under the 6 and 1' designation) or position 1 (under the 4 and ii designation) is thought to influence the nucleotide binding preference by the PPR motifs. The predicted RNA sequence generated by these combinations of three amino acids was very similar to what was predicted by the two-amino acid code. However at position 2, where a C or a U is predicted by the two-amino acid PPR code, an A or G is predicted with the three-amino acid code, which is what is actually observed at that position. Additionally, at position 11, where only a G is predicted by the two-amino acid PPR code, a G or a U is predicted by the three-amino acid PPR code, which is what is observed in all five cis-elements.

Interaction of QED1 with Other Known Editing Factors—Recently, protein factors other than PPR proteins have been found to be essential for editing in plant organelles. Members of the RIP/MORF protein family have been shown to be either plastid or mitochondrial editing factors, and one is active in both organelles (1, 4, 30). ORRM1, which belongs to a clade of RNA recognition motif (RRM) proteins, has been shown to be essential for editing of a subset of chloroplast C targets (3). We were therefore interested in determining whether QED1 interacts with any RIP family members or ORRM1 in yeast two-hybrid assays. QED1 was fused with the Gal4 DNA BD and co-expressed with members of the RIP family (RIP1, RIP2, RIP3, RIP5, and RIP9) as well as ORRM1 fused with the AD in yeast using growth on minimal medium (–histidine) as a reporter for protein–protein interactions. QED1 was also fused with the AD and co-expressed with BD-QED to test for possible homodimerization. Each gene was also co-expressed with the empty vector of the opposite vector type to ensure any yeast growth that was observed was not caused by auto-activation. Growth of yeast when BD-QED1 was co-expressed with each of the RIP family members indicated significant protein-protein interaction (Fig. 6). Unexpectedly, no protein-protein interaction was detected when BD-QED1 was co-expressed with AD-ORRM1 despite ORRM1 affecting the same sites as QED1 (Fig. 6A). Assays with BD-ORRM1 are uninformative because of autoactivation (data not shown). When BD-QED1 was co-expressed with AD-QED1, some growth was observed, indicating a weak protein-protein interaction (Fig. 6B).

A C-terminal Tag Interferes with QED1 Editing, whereas an N-terminal Tag Does Not—To produce a tagged QED1 for use in biochemical and complementation studies, we produced a construct with a 3×FLAG-StrepII tag fused to the C terminus of QED1, separated by a GGSGGSGS linker. When we transfected the C-tagged QED1 into qed1 protoplasts, we observed no complementation. We then moved the 3×FLAG-StrepII tag to the N terminus following a RecA chloroplast transit sequence and placed the GGSGGSGS linker between the tag and the PPR motifs. This construct was able to restore editing (Fig. 7). These results are consistent with reports that several mitochondrial PPR-DYW editing factors cannot tolerate C-terminal fusions to GFP nor to a His tag (31, 32). It is likely that
C-terminal tags interfere with the function of the C-terminal DYW motif in some PPR-DYW proteins.

Deletion of the E and/or DYW motifs of QED1 and RARE1 Results in Greatly Reduced Editing Efficiency—We examined the effect of deletions of the C-terminal region of QED1 on editing efficiency by comparing the restoration of editing efficiency in mutant protoplasts when the entire coding region of QED1 was transfected versus constructs deleted beginning with the E motif, the E/H11001 motif, or the DYW motif. Each of these QED1 constructs was cloned into an expression vector under a 35S promoter and expressed in qed1 protoplasts. A GFP marker that labels the endoplasmic reticulum was also expressed in the qed1 negative control to visualize transfection efficiency (~80%). The editing efficiencies of the five editing sites controlled by QED1 in each set of transfected protoplasts were measured using PPE (Fig. 8A). Wild-type Landsberg erecta protoplasts and qed1 protoplasts were transfected with GFP to control for transfection effects on editing efficiency. The full-length QED1 construct was able to complement the editing phenotype of the qed1 mutant at all five sites, restoring editing to about 40%, the lower efficiency than in wild type likely due to incomplete transfection or inadequate expression levels. Editing was almost undetectable when constructs with truncations encompassing...
Mutational Analysis of the DYW Domain of PPR Proteins

CRR22 snippet

Poter and DJW with

QED1 editing defect.

QED1 mutant protoplasts

QED1 protoplasts

The DYW motif in editing PPR-DYW factors shares some sequence and structural similarities to deaminases, including the conserved deaminase-like motifs (HSE, CXXCH) as denoted by the red lines along with the second putative zinc binding site and the PGC box designated by Hayes et al. (17).

either the E domain or the E+ domain were transfected, and only a very low level of editing was seen when the DYW deletion construct was transfected (Fig. 8A).

Several PPR-DYW editing factors had been reported to complement the editing defects of their respective mutants when their DYW domains were deleted (17, 33, 34). We therefore investigated whether RARE1, another PPR-DYW protein we had previously studied (16), could also tolerate a DYW deletion and restore editing to rare1 mutant protoplasts. Very little editing was detected when the RARE1 DYW truncation was transfected (Fig. 8A).

Further Analysis of the E and DYW Domains—The lack of complementation by truncated proteins allowed us to carry out mutagenesis of the E and DYW domains of QED1 and RARE1. The DYW motif in editing PPR-DYW factors shares some sequence and structural similarities to deaminases, including highly conserved residues that resemble the canonical HXE/CXXCH deaminase active sites (Fig. 9) (7, 35). To investigate whether these specific residues are important to the editing of C targets of QED1 and RARE1, we performed site-directed mutagenesis on the highly conserved HSE and CXXCH residues in the DYW motif of QED1 and in RARE1 (Figs. 10A and 11A).

For QED1, we created various constructs that targeted these conserved deaminase residues (Fig. 10A). In one construct, the HSE residues were all mutated to alanines, whereas in another construct, the five residues including the conserved cysteines (CDGCH) were all mutated to alanines. Because mutating five sequential residues could have an effect on the stability or folding of QED1, further constructs were created that targeted only the two cysteines.

Constructs were made in which the two cysteines were changed to two alanines or to two serines, the latter amino acid having greater structural similarity to cysteine than does alanine. Each of these mutated constructs was cloned into an expression vector under a 35S promoter and expressed in qed1 protoplasts along with full-length QED1 and ΔDYW for comparison. When we assessed editing at the five editing sites controlled by QED1, we observed that the protoplasts transfected by the mutated constructs exhibited little to no complementation of the editing defect. The QED1ΔDYW and QED1 with mutated HSE were able to support a greater level of editing than proteins in which the CXXCH signature was altered (Fig. 10B).

To investigate whether the cytidine deaminase signature of RARE1 was also essential, we performed site-directed mutagenesis on the DYW motif of RARE1. The HSE residues were again all mutated to alanines, whereas the targeted cysteines were mutated to serines (Fig. 10A). These mutated versions of RARE1 expressed under a 35S promoter in rare1 mutant protoplasts along with full-length RARE1 and ΔDYW for comparison. When editing at the five editing sites controlled by QED1, we observed that the protoplasts transfected by the mutated constructs exhibited little to no complementation of the editing defect. The QED1ΔDYW and QED1 with mutated HSE were able to support a greater level of editing than proteins in which the CXXCH signature was altered (Fig. 10B).

Two Terminal Cysteines in the DYW Motif of QED1 Are Important for Editing Function—We also investigated additional C-terminal regions of QED1 by site-directed mutagenesis. Hayes et al. (17) postulated that the deaminase-like residues CXXCH as well as CSC and a nearby His, four highly conserved residues near the C terminus of the DYW motif, bind the two zinc ions they detected in PPR-DYW factors EL1 and DYW1. Hayes et al. (17) also observed that a truncation of EL1 at the “PG box,” a 15-amino acid sequence spanning the boundary of
the E and E+ motif that is highly conserved in editing PPR-DYW, was unable to complement an ell1 mutant. The three terminal residues containing the two cysteines were mutated to alanines along with three well conserved residues of the PG box, PGC (Fig. 11A).

When these mutated versions of QED1 were expressed under a 35S promoter in qed1 protoplasts along with full-length QED1 and QED1/H9004 DYW for comparison, we found that editing was severely reduced when the terminal cysteines were

FIGURE 11. Mutagenesis of a second potential zinc binding site in the DYW motif of QED1 significantly reduces editing. A, residues in the E and DYW motif of QED1 were targeted for site-directed mutagenesis. Three residues (in red) in a conserved 15-amino acid sequence called the PGC box as defined by Hayes et al. (17) (in bold) and three residues (in red) near the end of the DYW motif of QED1 were mutated to alanines. B, results from the PPE assay in wild-type protoplasts (WT) and following transfection of GFP (qed1– or rare1–), the full-length un-mutagenized construct (Full), and mutagenized constructs targeting the PGC box (PGC or P) and the terminal cysteines (CC) into qed1 protoplasts. matK C640 PPE gel is shown. e, edited product; u, unedited product; o, oligonucleotide. g, genomic DNA. C, graphs of PPE data from rpoB C2432 and matK C640 sites are provided. The error bars indicate S.D.

the E and E+ motif that is highly conserved in editing PPR-DYW, was unable to complement an ell1 mutant. The three terminal residues containing the two cysteines were mutated to alanines along with three well conserved residues of the PG box, PGC (Fig. 11A).

When these mutated versions of QED1 were expressed under a 35S promoter in qed1 protoplasts along with full-length QED1 and QED1ΔDYW for comparison, we found that editing was severely reduced when the terminal cysteines were

FIGURE 10. Mutagenesis of the cytidine deaminase signature residues in QED1 and RARE1 significantly reduces editing. A, the deaminase residues in the DYW motif (designated in red) of QED1 were targeted using site-directed mutagenesis. Changes that were made are indicated above the targeted residues. B, PPE assay of editing in qed1 protoplasts following two separate transfections of the same batch of protoplasts with full-length (Full), GFP (qed1–), DYW-truncated QED1 (ΔDYW), and mutagenized versions of QED1 (CXXCH, HSE) when compared with wild-type protoplasts (WT(+)); g, genomic DNA; e, edited product; u, unedited product; o, oligonucleotide. C, PPE assay of editing in qed1 protoplasts using a different batch of protoplasts than in B and additional mutagenized versions of QED1 (AXXAH, SXSSH). A PPE gel for matK C640 is shown; graphs of editing quantifications for matK C640 and rps12 C58 PPE gels are provided. D, PPE assay of editing in the same batch of rare1 protoplasts following two separate transfections with full-length RARE1 (Full), GFP (rare1–), DYW-truncated RARE1 (ΔDYW), and mutated versions SXSSH and HSE. The error bars indicate S.D.
Mutational Analysis of the DYW Domain of PPR Proteins

Mutational Analysis of the DYW Domain of PPR Proteins

FIGURE 12. Expression of full-length QED1 and mutagenized QED1 (SXXSH) in mutant protoplasts and effect on editing of three C targets. A, immunoblot probed with α-FLAG antibody. qed1 protoplasts were transfected with a construct expressing either the entire wild-type QED1 coding region (WT) or the mutant QED1 (SXXSH), tagged on their N termini with a 3 × FLAG-Strepl II tag. Protein was extracted from the transfected protoplasts, from un-transfected wild-type protoplasts (WT) and qed1 protoplasts transfected with a GFP gene (qed1). 4 μg of protein from each sample was run on an SDS-PAGE gel. 10 μg of FLAG-BAP (bacterial alkaline phosphatase) was included as a positive control. Mature QED1 plus the 3 × FLAG-Strepl II tag is predicted to be ~80 kDa in size. B, editing quantifications for ndhB C872, rpoB C2432, and matK C640 from RNA extracted from the same protoplast samples are shown underneath. The error bars indicate S.D.

mutated (Fig. 11B). Editing was only slightly reduced when PGC was altered to AAA (Fig. 11B).

QED1 Mutated through Site-directed Mutagenesis Is Stably Expressed in Protoplasts—A lack of complementation by a construct carrying a mutated coding region could occur if the mutant protein is highly unstable. Although it seems unlikely that multiple constructs targeting different residues could all fail to accumulate, we examined expression of two constructs by tagging them with a 3 × FLAG-Strepl II tag at the N terminus (Fig. 12). The mature wild-type QED1 and QED1 (CXXCH mutated to SXXSH) were tagged at their N termini with a 3 × FLAG-Strepl II tag as well as a RecA chloroplast targeting sequence and a GGSGGGS linker. These N-tagged constructs were expressed from a 35S promoter in qed1 mutant protoplasts. RNA was extracted from a portion of the protoplasts, and the remainder of the cells was lysed to obtain protein for immunoblotting. The protein extracts from each set of transfected protoplasts, as well as protein extract from untransfected wild-type and qed1 protoplasts and along with bacteria alkaline phosphatase (BAP) tagged with FLAG (Sigma), were then probed with α-FLAG antibody to probe for expression of the N-tagged constructs (Fig. 12). As expected, no signal was observed in the wild-type and qed1 protein extracts from protoplasts that were not transfected by the FLAG-tagged constructs. In each of the samples of protoplasts that were transfected by the FLAG-tagged versions of QED1, we observed a band of about 80 kDa, which is the expected size of mature QED1 with the added FLAG-Strepl II tag (Fig. 12).

DISCUSSION

Until this study, no chloroplast PPR protein had been found to control more than three editing sites. We have found that QED1 is responsible for the editing of five different C targets in five different chloroplast genes, accD, matK, ndhB, rpoB, and rps12. Three of these sites (matK, ndhB, and rpoB) are found in the coding region of the gene, whereas the two other sites affect non-coding regions. A PPR protein had been assigned as a site-specific editing factor only to the site within the first intron of rps12. While this work was in progress, another group published RT-PCR bulk sequencing assays of mutants containing only the weak qed1 allele from the Salk collection in the gene they named OTP81, and they detected a deficiency only in rps12 intron editing (15). They reported that failure to edit this site did not affect splicing of rps12 (15).

Our truncations of the E or E+ motif resulted in a complete loss of editing activity, a result consistent with other studies in which an editing PPR protein was truncated before the E or E+ motif (17, 33, 37). The E/E+ motif is an essential part of the RNA editing reaction, although its role is unclear. Although there is higher conservation between the DYW motifs of editing factors, the E/E+ motif still shows a higher degree of conservation when compared with the PPR motifs. The E/E+ motif also appears to contain two particularly degenerate PPR motifs that could serve in an RNA or protein binding role. If the DYW motif is the deaminase, perhaps the E/E+ motif plays a role in recruiting or mediating protein-protein interactions between the factors that provide the RNA editing activity, which may include DYW motifs that act in trans. E motifs swapped between two mitochondrial and two chloroplast editing factors could not restore editing at those specific sites, but E domains exchanged with an E domain from a protein targeted to the same organelle did give a functional chimeric protein (38), implying that there may be some organelle specialization in the function of these motifs.

PPR-DYW proteins undoubtedly interact with other components of the protein complexes that carry out plant organelle RNA editing. Recent studies of non-PPR protein editing factors have revealed that the PPR-DYW site recognition factors are not the only proteins that act site-specifically to mediate editing. For example, RIP1, RIP2, RIP9, and ORRM1 are each essential for editing of particular chloroplast C targets, but dispensable for others (1, 3, 30). However, mutation in RIP2, RIP9, or ORRM1 affected the efficiency of editing of the five sites controlled by QED1, suggesting that the RIP and ORRM1 proteins may act together with QED1 within a subset of editosomes. As would be predicted from such a model, RIP1, RIP2, and RIP9 all interacted with QED1 in our yeast two-hybrid assays. Unexpectedly, no interaction was detected between QED1 and
Mutational Analysis of the DYW Domain of PPR Proteins

ORRM1. One possible explanation is that QED1, plastid RIP proteins, and ORRM1 are in the same complex, but QED1 directly associates with RIP proteins but not with ORRM1, which may instead interact only with one or more RIP proteins. It is also possible that QED1 does interact with ORRM1 in chloroplast, but not in yeast. If QED1/ORRM1 interaction requires the presence of the chloroplast RNA transcript carrying the cis-element to which QED1 binds, the interaction will not occur in yeast where the transcript does not exist. However, ORRM1 has been shown to directly interact with other editing PPR proteins in yeast, so this latter explanation is less likely.

The chloroplast factor QED1 also interacts with two mitochondrial RIP proteins: RIP3 and RIP5. A similar promiscuous interaction between PPR proteins and RIP protein was also reported by Takenaka et al. (4) in which mitochondrial PPR editing factors MEF1, MEF9, and MEF21 interacted with the plastid RIP proteins RIP2 and RIP9 (also named MORF2 and MORF9). Such interactions are unlikely to occur in planta because the PPR proteins are compartmentalized in different organelles.

QED1 exhibited weak interaction with itself in the yeast two-hybrid assay. A PPR protein, named PPR10, was reported to form an antiparallel homodimer in RNA-free crystals (39). However, the authors also reported that binding of RNA weakened the homodimerization. Another PPR protein, THA8, was also reported to form an asymmetric homodimer, but unlike PPR10, RNA binding induced dimerization (40). As yeast does not contain the target chloroplast RNA with which QED1 normally interacts, it is possible that no dimerization of QED1 occurs in vivo. However, the presence of the target chloroplast RNA could also strengthen the interaction.

Salone et al. (7) hypothesized that the DYW motif served as the deaminase in the RNA editing reaction because it contains a set of residues that are not only highly conserved between DYW motifs, but are also highly conserved with residues that comprise the active site in known deaminases. The predicted DYW protein structure also correlates with the actual or predicted structure of other known deaminases (35). Furthermore, evolutionary analysis indicates that the presence of the DYW motif correlates with the presence of C to U editing in plants (7, 41–43). An issue regarding this hypothesis is the observation that the DYW motif is not needed for the editing function of four PPR-DYW editing factors (CRR22, CRR28, OTP82, and ELI1) (17, 33, 34), and there are editing factors with an E motif that was also missing in CRR4 (45). The mitoochondrional RIP protein MEF1 also requires its DYW motif to restore editing to mef1 protoplasts (32).

We have discovered two more editing factors, QED1 and RARE1, which not only require the DYW motif for efficient function, but specifically, the conserved deaminase residues. In some cases, QED1 lacking its DYW motif conferred a greater amount of editing than samples with site-specific mutations within the motif. This finding might be due to another DYW motif-containing protein being able to act in trans at low efficiency when the DYW motif of QED1 is absent. We have verified that some of the mutant proteins accumulate in chloroplasts, so the lack of editing activity cannot be ascribed to the absence of the protein. When we mutated the cysteines in the putative second zinc binding site described by Hayes et al. (17), we observed an effect on RNA editing similar to what was observed when we mutated the cysteines in the deaminase motif. We also observed that the placement of a 3×FLAG-StrepII tag at the C terminus of QED1 greatly interferes with RNA editing function, whereas a 3×FLAG-StrepII tag at the N terminus does not, which also indicates that the function of the C-terminal DYW motif is critical for editing activity.

Although we have identified PPR-DYW proteins that require their DYW domain for editing activity, it is still unclear why other PPR-DYW proteins, such as CRR22 and CRR28, can function without the DYW domain. The composition of editosomes is diverse, with different members of other gene families present in editosomes with certain PPR proteins but not with others. Until more is known about the complete composition of the editosome and its molecular structure, it will not be possible to understand why the DYW domain is required in some PPR proteins but not in others.

While this manuscript was in preparation, an article by Boussardon et al. (46) appeared in which zinc binding and the deaminase signature of DYW1 were analyzed by mutagenesis. In contrast to the work by Hayes et al. (17) that described binding of two zinc atoms by PPR-DYW editing factors, ELI1 and DYW1, binding of only one zinc by a GST fusion with DYW1 was reported by Boussardon et al. (46), which is at odds with our finding that the putative second zinc binding site is important for RNA editing in QED1 and the previously published results of Hayes et al. (17). Unlike the proteins we have analyzed, DYW1 lacks PPR motifs, but apparently works in conjunction with the PPR protein CRR4 to mediate editing to specify an AUG start codon in the chloroplast ndhD transcript. The HXE and CXXC residues conserved in DYW1 were mutated, resulting in a loss of editing (46), which is consistent with our finding that these residues characteristic of cytidine deaminases also present in PPR-DYW factors are essential for chloroplast RNA editing.

REFERENCES

1. Bentolila, S., Heller, W. P., Sun, T., Babina, A. M., Friso, G., van Wijk, K. J., and Hanson, M. R. (2012) RIP1, a member of an Arabidopsis protein family, interacts with the protein RARE1 and broadly affects RNA editing. Proc. Natl. Acad. Sci. U.S.A. 109, E1453–E1461

2. Kotera, E., Tasaka, M., and Shikanai, T. (2005) A pentatricopeptide repeat protein is essential for RNA editing in chloroplasts. Nature 433, 326–330

3. Sun, T., Germain, A., Giloteaux, L., Hammani, K., Barkan, A., Hanson, M. R., and Bentolila, S. (2013) An RNA recognition motif-containing protein is required for plastid RNA editing in Arabidopsis and maize. Proc. Natl. Acad. Sci. U.S.A. 110, E1169–E1178

4. Takenaka, M., Zehrmann, A., Verbitskiy, D., Kugelman, M., Härtel, B., et al. (2013) An RNA recognition motif-containing protein with DYW motif is critical for editing activity. Mutational Analysis of the DYW Domain of PPR Proteins

5. Takenaka, M., Salone, A., Verbitskiy, D., Kugelman, M., Härtel, B., et al. (2013) An RNA recognition motif-containing protein with DYW motif is critical for editing activity. Mutational Analysis of the DYW Domain of PPR Proteins
Mutational Analysis of the DYW Domain of PPR Proteins

22. Herkel, J., Siemerling, K. R., Prashe, D. C., and Hodge, S. (1997) Removal of a cryptic intron and subcellular localization of green fluorescent protein are required to mark transgenic Arabidopsis plants brightly. Proc. Natl. Acad. Sci. U.S.A. 94, 2122–2127

23. Bentolila, S., Elliot, L. E., and Hanson, M. R. (2008) Genetic architecture of mitochondrial editing in Arabidopsis thaliana. Genetics 178, 1693–1708

24. Peeters, N. M., and Hanson, M. R. (2002) Transcript abundance supersedes editing efficiency as a factor in developmental variation of chloroplast gene expression. RNA 8, 497–511

25. Waterhouse, A. M., Proctor, J. B., Martin, D. M., Clamp, M., and Barton, G. J. (2009) Jalview Version 2: a multiple sequence alignment editor and analysis workbench. Bioinformatics 25, 1189–1191

26. Hildebrand, M., Hallick, R. B., Passavant, C. W., and Bourque, D. P. (1988) Trans-splicing in chloroplasts: the rps12 loci of Nicotiana tabacum. Proc. Natl. Acad. Sci. U.S.A. 85, 372–376

27. Koller, B., Fromm, H., Galun, E., and Edelmann, M. (1987) Evidence for in vivo trans-splicing of pre-mRNAs in tobacco chloroplasts. Cell 48, 111–119

28. Bentolila, S., Oh, J., Hanson, M. R., and Bukowski, R. (2013) Comprehensive high-resolution analysis of the role of an Arabidopsis gene family in RNA editing. PLoS Genet, e1003584

29. Brehme, N., Zehrmann, A., Verbitskiy, D., Härter, B., and Takenaka, M. (2014) Mitochondrial RNA editing PPR proteins can tolerate protein tags at E as well as at DYW domain termini. Front. Plant Sci. 5, 127

30. Zehrmann, A., Verbitskiy, D., Härter, B., Brennicke, A., and Takenaka, M. (2010) RNA editing competence of trans-factor MEF1 is modulated by ecotype-specific differences but requires the DYW domain. FEBS Lett. 584, 4181–4186

31. Okuda, K., Chateigner-Boutin, A. L., Nakamura, T., Delannoy, E., Sugita, M., Myougou, F., Motohashi, R., Shinozaki, K., Small, I., and Shikanai, T. (2009) Pentatricopeptide repeat proteins with the DYW motif have distinct molecular functions in RNA editing and RNA cleavage in Arabidopsis chloroplasts. Plant Cell 21, 146–156

32. Okuda, K., Hammani, K., Tansz, S. K., Peng, L., Fukao, Y., Myougou, F., Motohashi, R., Shinozaki, K., Small, I., and Shikanai, T. (2010) The pentatricopeptide repeat protein OTP82 is required for RNA editing of plastid ndhB and ndhG transcripts. Plant J. 61, 339–349

33. Iyer, L. M., Zhang, D., Rogozin, I. B., and Aravind, L. (2011) Evolution of the deaminase fold and multiple origins of eukaryotic editing and mutagenic nucleic acid deaminases from bacterial toxin systems. Nucleic Acids Res. 39, 9473–9497

34. Taly, J. F., Magis, C., Bussotti, G., Chang, J. M., Di Tommaso, P., Erb, I., Espinosa-Carrasco, J., Kemena, C., and Notredame, C. (2011) Using the T-Coffee package to build multiple sequence alignments of protein, RNA, DNA sequences and 3D structures. Nat. Protoc. 6, 1669–1682

35. Okuda, K., Myougou, F., Motohashi, R., Shinozaki, K., and Shikanai, T. (2007) Conserved domain structure of pentatricopeptide repeat proteins involved in chloroplast RNA editing. Proc. Natl. Acad. Sci. U.S.A. 104, 8178–8183

36. Chateigner-Boutin, A.-L., Colas des Francs-Small, C., Fujiy, S., Okuda, K., Tansz, S. K., and Small, I. (2013) The E domains of pentatricopeptide repeat proteins from different organelles are not functionally equivalent for RNA editing. Plant J. 74, 935–945

37. Yin, P., Li, Q., Yan, C., Liu, Y., Liu, J., Yu, F., Wang, Z., Long, J., He, J., and Yang, H. W. (2013) Structural basis for the modular recognition of single-stranded RNA by PPR proteins. Nature 504, 168–171

38. Ke, J., Chen, R. Z., Ban, T., Zhou, X. E., Gu, X., Tan, M. H., Chen, C., Kang, Y., Brunzel, J. S., Zhu, J. K., Melcher, K., and Xu, H. E. (2013) Structural basis for RNA recognition by a dimeric PPR-protein complex. Nat. Struct. Mol. Biol. 20, 1377–1382

39. Rüdinger, M., Polskakiewicz, M., and Knoop, V. (2008) Organellar RNA editing and plant-specific extensions of pentatricopeptide repeat proteins in jungermannioid but not in marchantid liverworts. Mol. Biol. Evol. 25, 1405–1414

40. Knoop, V., and Rüdinger, M. (2010) DYW-type PPR proteins in a heteroloobosean protist: plant RNA editing factors involved in an ancient horizontal gene transfer? FEBS Lett. 584, 4287–4291


Mutational Analysis of the DYW Domain of PPR Proteins

43. Rüdinger, M., Volkmar, U., Lenz, H., Groth-Malonek, M., and Knoop, V. (2012) Nuclear DYW-type PPR gene families diversify with increasing RNA editing frequencies in liverwort and moss mitochondria. *J. Mol. Evol.* **74**, 37–51

44. Takenaka, M., Verbitskiy, D., Zehrmann, A., and Brennicke, A. (2010) Reverse genetic screening identifies five E-class PPR proteins involved in RNA editing in mitochondria of *Arabidopsis thaliana*. *J. Biol. Chem.* **285**, 27122–27129

45. Boussardon, C., Salone, V., Avon, A., Berthomé, R., Hammani, K., Okuda, K., Shikanai, T., Small, I., and Lurin, C. (2012) Two interacting proteins are necessary for the editing of the ndhD-1 site in *Arabidopsis* plastids. *Plant Cell* **24**, 3684–3694

46. Boussardon, C., Avon, A., Kindgren, P., Bond, C. S., Challenor, M., Lurin, C., and Small, I. (2014) The cytidine deaminase signature HxEx(x)n CxxC of DYW1 binds zinc and is necessary for RNA editing of ndhD-1. *New Phytol.* **203**, 1090–1095