DYW domain structures imply an unusual regulation principle in plant organellar RNA editing catalysis

Mizuki Takenaka, Sachi Takenaka, Tatjana Barthel, Brody Frink, Sascha Haag, Daniil Verbitskiy, Bastian Oldenkott, Mareike Schallenberg-Rüdinger, Christian G. Feiler, Manfred S. Weiss, Gottfried J. Palm and Gert Weber

RNA editosomes selectively deaminate cytidines to uridines in plant organellar transcripts—mostly to restore protein functionality and consequently facilitate mitochondrial and chloroplast function. The RNA editosomal pentatricopeptide repeat proteins serve target RNA recognition, whereas the intensively studied DYW domain elicits catalysis. Here we present structures and functional data of a DYW domain in an inactive ground state and activated. DYW domains harbour a cytidine deaminase fold and a C-terminal DYW motif, with catalytic and structural zinc atoms, respectively. A conserved gating domain within the deaminase fold regulates the active site sterically and mechanistically in a process that we termed gated zinc shutter. Based on the structures, an autoinhibited ground state and its activation are cross-validated by RNA editing assays and differential scanning fluorimetry. We anticipate that, in vivo, the framework of an active plant RNA editosome triggers the release of DYW autoinhibition to ensure a controlled and coordinated cytidine deamination playing a key role in mitochondrial and chloroplast homeostasis.

PLACE

Plant RNA editing specifically converts several hundreds of cytidines to uridines in mitochondrial and chloroplast transcripts. The RNA editing activity has to be stringently suppressed in the cytosol and stimulated only in the organelles at the target site, as no C-to-U RNA editing has been described in nuclear transcripts in plants. Nuclear-encoded pentatricopeptide repeat (PPR) proteins with a C-terminal DYW domain have been characterized as site-specific factors for C-to-U RNA editing in plant mitochondria and plastids. RNA substrate recognition is conferred by the PPR tract, whereas the exact role of the DYW domain, which can be also recruited to an editing site in trans, has not been clarified. The DYW domain, which was named by the highly conserved last three amino acids, aspartate, tyrosine and tryptophan, has been proposed as the best candidate to elicit deamination employing a HxE(x)CxxC zinc ion binding signature. Indeed, DYW domains from DYW1 and ELI1 demonstrated zinc ion binding capacity, and recent orthogonal E.coli as well as in vitro experiments with a single DYW domain elicits catalysis. Apart from a cytidine deaminase fold, DYW domains contain a characteristic DYW motif, stabilized by a zinc atom, as well as a gating domain that controls zinc-mediated catalysis sterically and catalytically. The catalytic regulation hallmarks an unusual protein regulation principle where, upon activation, a major movement of the gating domain alters the coordination around the catalytic zinc atom while in the inactive state, the zinc is inhibited by its coordination setting. We employed in vivo RNA editing assays to map the potential RNA path on the DYW domain and identify key residues required for regulation and catalysis to occur. Finally, RNA in vitro editing and thermal shift assays consolidate the structural data and confirm a tetrahydrouridine or nucleotide triphosphate-triggered activation mirroring the two different conformational states. Beyond the identification of an unusual principle in metalloenzyme regulation, our results reveal key mechanisms in plant organellar RNA editing catalysis, its autoinhibition and have far-reaching implications for mitochondrial and chloroplast homeostasis.

RESULTS

Crystal structure of the Arabidopsis thaliana OTP86<sup><sub>11,12</sub></sup>, Here we report crystal structures of the DYW domain of an Arabidopsis thaliana (A. thaliana) plastid RNA editing factor, OTP86, as the outcome of a solubility and crystallization screening of over 100 different DYW domain constructs from 30 PPR proteins. OTP86 was characterized as a site-specific factor for an editing site in rps14 transcripts. The protein consists of 20 N-terminal PPR repeats, E1 and E2 motifs, which are predicted to have a PPR-
tetrameric peptide-repeat-like (TPR-like) fold and a C-terminal DYW domain.31

To initially assess whether the OTP86 DYW domain (OTP86DYW) is an active editing factor, we conducted in vivo orthogonal RNA editing assays in *Escherichia coli* (*E. coli*) and in vitro assays with purified proteins. Both methods verified the cytidine deaminase activity of the OTP86 DYW domain when fused with the PPR tract of the moss *Physcomitrella* Pp56 protein (Supplementary Fig. 1a–c).22,23 When the catalytically important E894 of OTP86DYW was replaced by an alanine, editing was abolished.

We then set out to pioneer the structural characterization of DYW domains exemplified by OTP86DYW. Several years of crystallization attempts were severely hampered by the very limited amounts of soluble OTP86DYW (residues G826 to W960), which migrates at a molecular weight of about 15kDa in size-exclusion chromatography, indicating a monomeric state (Supplementary Fig. 2a).

Finally, we obtained crystals of OTP86DYW belonging to space group C2 and diffracting to a resolution of 2.5 Å (Supplementary Fig. 3a and Supplementary Table 1). The structure was solved by single-wavelength anomalous dispersion (SAD) phasing harnessing four zinc atoms (see Methods, Supplementary Table 1 and Supplementary Fig. 3b,c for details).

The fold of OTP86DYW is highly similar to cytidine deaminases but has prominent additional features (Fig. 1a–e). A comparison of OTP86DYW with *E. coli* cytidine deaminase (PDB ID: 1CTU; ref. 32) reveals an overall similarity (r.m.s.d. = 2.4 Å for 72 of 132 residues superimposed) to the typical core deaminase fold comprising five β-strands flanked by two α-helices.33,34 The region previously termed PG box covers the first two β-strands of the deaminase domain (Fig. 1a–c).22,23 Remarkably, the deaminase fold of OTP86DYW is interrupted by an insertion of about 55 residues that bridge β-strand 2 and α-helix 2 (Figs. 1b and 2). The insertion is composed of an amphipathic α-helix that runs across one face of the entire structure contacting both α-helices of the deaminase fold with conserved hydrophobic residues (Supplementary Fig. 3d and Fig. 2) and re-enters the deaminase fold via a highly conserved β-finger at α-helix 2, which in turn harbours the HxE(x),CxxC motif, crucial to catalysis and substrate binding21,22,23. This motif has a high similarity to the cytidine deaminase signature HxE(x),PxCxxC and contains a catalytically important glutamate residue (E894 in OTP86) only the proline is not conserved in DYW domains (Fig. 2)35. Contrasting the large inserted domain of OTP86DYW, *E. coli* cytidine deaminase only contains a smaller loop which instead points away from the active site permitting nucleotide entry (Fig. 1a). We conclude that the OTP86DYW active site seems to have limited accessibility for sub-

active site permitting nucleotide entry (Fig. 1a). We conclude that

When employing the coordinates of OTP86DYW, structure solution revealed a region of positively charged residues spanning across the active site and passing in between the base of the gating domain's β-finger and the DYW motif. As RNA bases around the editing site are not conserved, this probably represents the path of the negatively charged RNA backbone, which is placed for catalysis by the PPR tract after or concomitant to activation of the DYW domain (Fig. 1c).

Crystal structure of an activated *A. thaliana* OTP86DYW. As crystal soaking experiments with substrate, product or different short RNA trinucleotides were unsuccessful, we attempted co-crystallization of OTP86DYW with the well-characterized deam-

inase inhibitor tetrahydrouridine (THU). Along this approach, we observed several new crystallization conditions that indicated a different crystallization behaviour due to the presence of THU. Finally, we obtained crystals of OTP86DYW with space group P21 21 2, which diffracted to a resolution of 1.65 Å (Supplementary Table 1). When employing the coordinates of OTP86DYW, structure solution by molecular replacement failed; however, four copies of a trun-
cated model missing the gating domain were successfully placed in the asymmetric unit (see Methods, Supplementary Table 1 and Supplementary Fig. 5). Explaining the failed molecular replacement, OTP86DYW had clearly changed its conformation substantially in the presence of THU towards an activated (OTP86DYW*) state (Fig. 3a, Supplementary Fig. 6 and Supplementary Video 1). The conformational change mainly involved the β-hairpin of the gating domain (which now adopts an extended β-strand conformation), and its connection to α-helix 1 (gating domain) and α-helix 2 (deaminase domain). It is widely accepted that confor-
mational switches of β-fingers may take part in the regulation of macromolecular complexes as observed for the RNase H domain in the spliceosomal Prp8 protein35,36. The conformational change has a marked effect on the active site architecture, in particular zinc coordination. The inactive structure zinc coordination is maintained by coordinating H892, C920, C923 and a more distant water molecule, whereas the catalytically important E894 is ionically bonded to K915 (Fig. 3b); K915 also hydrogen bonds to S828 and S893. In this configuration, the ion pair will reduce the basic character of E894 and hinder the required deprotonation of the deaminating water, which is not productively coordinated by Zn1 and also not contacted by E894. We reason that beyond the steric inhibition through the gating domain, K915 has to be released before efficient catalysis can occur. This notion is corroborated by the activated OTP86DYW* (Fig. 3d), in which K915 points away from E894. Although E894 is conserved in all deaminases, K915 is restricted to DYW domains (Figs. 1c and 2)35. The conformational changes upon activation involve several larger backbone torsion angle movements of the gating domain’s β-hairpin.
The direct effects of the gating domain’s conformational change on OTP86\textsuperscript{DYW} catalytic activation via Zn1 coordination are evident from the detailed structural comparison of the zinc coordination and E894 (Fig. 3d,e, and Supplementary Videos 1 and 2). Remarkably, the conformational change of H892 from the main chain dihedral angles $\phi/\psi = 58^\circ/42^\circ$ (inactive) to $\phi/\psi = -74^\circ/152^\circ$ (active) and a concomitant repositioning of its side chain elicits a pervasive impact on the Zn1 coordination geometry (Fig. 3b,c). When superimposing residues C920–C923 of both structures, activation moves the coordinating nitrogen of H892 by 2 Å, with a concomitant rotation around the Zn1 coordination sphere by about 35° that harnesses the zinc ligands C920 and C923 as a rotation axis; C920, C923 and Zn1 remain largely unaffected during activation. The restructuring of the active site reduces the zinc–water/water–E894 distances from 3.07 Å/3.93 Å (inactive conformation) to 2.15 Å/2.53 Å, thereby activating the mechanistically important water molecule (Supplementary Video 2). The altered H892 positioning permits the remotely located water molecule to be attracted to Zn1 as a fourth coordination ligand poised for the deamination reaction (Fig. 3d,e and Supplementary Figs. 5 and 6). In OTP86\textsuperscript{DYW\textsuperscript{W}}, distances and angles of the zinc ligands are in agreement with a catalytically competent reaction centre\textsuperscript{41,42}. Furthermore, the strand length of the $\beta$-finger is extended upon activation, still maintaining the original backbone hydrogen bonding residue pairs of the inactive OTP86\textsuperscript{DYW}. The side chain of H890, which shields the active site as a counterpart relative to the zinc coordination sphere (formed by H892, C920 and C923) in the inactive OTP86\textsuperscript{DYW\textsuperscript{W}}, is repositioned far away from the active site by about 13 Å in OTP86\textsuperscript{DYW\textsuperscript{W}} (compare Fig. 3b,c, Supplementary Video 1). Unexpectedly, THU could not be located in the electron density, which implies a crucial role in triggering activation but not as tightly bound inhibitor. We are not aware of a comparable mechanism and thus coin this catalytic activation mechanism of DYW domains, and probably other metalloenzymes, gated zinc shutter.

Structural comparison of the OTP86 DYW domain to other cytidine deaminases. A comparison of OTP86\textsuperscript{DYW} with known ligand-bound deaminase domains confirms the presence of a complete active site for catalysis and fortifies the notion of a steric autoregulatory mechanism for DYW domains (Fig. 4a–i). When comparing OTP86\textsuperscript{DYW} with cytidine deaminase from mouse bound to cytidine (MmCD), or human APOBEC3A in complex with a short DNA (HsAPOBEC3A), nearly all of the residues required for nucleotide binding are present in OTP86\textsuperscript{DYW} and located at corresponding positions (Fig. 3b–d,g)\textsuperscript{43,44}. For example, all atoms of

**Fig. 1 | Crystal structure of the A. thaliana OTP86 DYW domain.** a, Superimposition of OTP86\textsuperscript{DYW} (marine) with E. coli cytidine deaminase (EcCD\textsuperscript{2}, cyan) bound to the inhibitor zebularine (not shown) (PDB-ID: 1CTU; ref. 33). The consensus deaminase zinc ions are shown as green (OTP86\textsuperscript{DYW}, Zn1) and light-green (EcCD\textsuperscript{2}, Zn2) spheres, a zinc ion partially coordinated by the DYW motif is shown in yellow (Zn2). b, The OTP86\textsuperscript{DYW} structure defines a paradigmatic organization for DYW domains. The cytidine deaminase domain (slate) coordinates a zinc ion (Zn1, green) three-fold with H892, C920 and C923, the fourth position is occupied by a water molecule (W, white sphere). The deaminase domain is interrupted by a gating domain (orange) and terminates with a DYW motif (red), partially coordinating a second zinc ion (Zn2, yellow). c, A close-up view on the cytidine deaminase active site, with catalytically relevant residues shown as sticks. d, A close-up view of the DYW motif and the flanking $\beta$-strand 7 as well as $\alpha$-helix 3. e, Electrostatic surface potentials as indicated by the colour scale bar (bottom), obtained by APBS version 1.5 and plotted on the surface of OTP86\textsuperscript{DYW}. Residues involved in zinc coordination are shown as sticks; zinc atoms are as in b. Rotation symbols indicate the views relative to b. Interacting residues are shown as sticks and coloured by atom type. Blue, nitrogen; red, oxygen; yellow, sulfur; carbons take the colour of the respective molecule. Dashed lines represent hydrogen bonds, whereas thick grey dashed lines indicate zinc coordination. Dashed lines in the ribbon plots represent residues 842–844 not clearly defined by electron density.
residues coordinating Zn1 and E894, as well as the backbone R918 carbonyl, C920 and S893 amides (contacting the respective base in MmCD and HsAPOBEC3A) of OTP86, superimpose to their mouse and human equivalents with r.m.s.d. values of 1.0 and 0.9 Å, respectively. The backbone carbonyl oxygen of R918 or the backbone amide of C920 are within hydrogen bonding distance to the amine of the base or the activated water molecule, respectively (compare Fig. 4b,c with Fig. 4d,g). Likewise, the backbone carbonyl oxygen of S893 (OTP86) may contact the keto group of the bound cytidine as for A66 in MmCD or A71 in HsAPOBEC3A. L197 of OTP86GW (Fig. 4b,c) has equivalent residues (I87 in MmCD or W98 in human APOBEC3A) that stack on the edited base (Fig. 4b,d,g). H70 in HsAPOBEC3A adopts a similar side chain conformation as H892, however, only in OTP86GW*, implying a role in base stacking upon activation (compare Fig. 4b,c,g). The OTP86-equivalent residue for MmCD N54 or HsAPOBEC3A N57, both of which contact the sugar 3’ oxygen, could not be identified. This residue may also be part of a region preceding the OTP86GW* deaminase fold (or PG box), which is missing in our structure. Hence we conclude that the OTP86GW* active site and the positioning of the base targeted for deamination is nearly identical to other cytidine deaminases.

The absence of the region preceding the PG box from our crystallization constructs may have impeded our attempts to obtain structures of OTP86GW* bound to substrate-related molecules.
A superimposition of bound nucleotides of known deaminase structures has further implications for the OTP86 activation mechanism. To investigate the nucleotide binding mode of OTP86DYW*, we superimposed active site residues of *Mm*CD and *Hs*APOBEC3A onto OTP86DYW in its activated and inactive states and compared the substrate positions (Fig. 4e,f,h,i). For example, the cytidine bound to *Mm*CD causes steric clashes with the β-finger of the OTP86DYW gating domain when positioned in the active site of OTP86DYW. Due to the conformational change of the gating domain upon OTP86DYW activation, this inhibition is released (compare Fig. 4e,f,i). When comparing OTP86 structures with DNA-bound human APOBEC3A, the +1 nucleotide (3’ of the active site) causes steric clashes with the gating domain only in the inactive OTP86DYW state, but not in OTP86DYW*. In conclusion, several superimposed substrate nucleotides suggest a steric inhibition by the OTP86 gating domain and probably other DYW deaminases as key residues (1) for nucleotide positioning and (2) participating in the conformational changes show a high degree of conservation (Fig. 2). These observations consolidate the notion of an autoinhibited ground state of OTP86DYW, which is paradigmatic for all PPR proteins with a DYW domain.

**In vivo RNA editing assays with OTP86DYW and variants.** To cross-validate the structural data and also probe the DYW domain surface, we conducted orthogonal in vivo RNA editing assays in *E. coli* employing PPR56PRE12-OTP86DYW mutants (Fig. 5). The solubility of the mutants was assessed by a western blot employing the soluble fraction of the respective cell lysates (Supplementary Fig. 7). To this end, the reduced activities of for example, L856, R912, T914, D922 can be explained by the very limited solubility of the respective fusion proteins. By contrast, K555A (PPR56 numbering, corresponding to position K823 in OTP86) is soluble and the mutant has a dramatically reduced editing activity. In OTP86, the equivalent lysine is located directly before the PG box at position 823 and may contact the sugar of the edited nucleotide or the acidic phosphate backbone, for example as N54 in *Mm*CD (see Fig. 4d). L889 directly precedes the active site as part of the β-finger, changes its position upon activation and may contact the RNA substrate remotely from the edited base, probably explaining its reduced in vivo RNA editing activity (Fig. 3b,c). H892 is a key regulatory residue as it alters its zinc coordination position upon activation, which poises the active site for the reaction. An alanine at this position is inactive as it is not suitable as a zinc coordination ligand. A cysteine may coordinate the zinc; however, H892C is also inactive. We reason that either the cysteine side chain does not provide the necessary flexibility to undergo a dramatic repositioning as that of histidine does. Alternatively, cysteine is a strong coordination ligand of zinc compared with histidine and may thus reduce zinc reactivity. S828 and S893 apparently play an important inhibitory role when contacting K915 and tethering it to the catalytically important E894 (Fig. 3b and Supplementary Note 3).

Mutants of the catalytic residue E894 retain their solubility upon mutation to alanine or the structurally analogous uncharged glutamine; however, both mutants are inactive, which consolidates their important role in deamination catalysis.32,33,55. Albeit soluble to a low degree (compare with D922), the R895A mutant is probably inactive due to a structural destabilization of the active site. Alternatively, R895 may be crucial to catalysis according to the previously described zinc charge compensation.54 Interestingly, the hydrogen bond donor of R895 changes from the terminal N° to the weaker bridging N° during activation. In the active conformation, R895 hydrogen bonds to D872 (Fig. 3b,c). This stabilizing effect is missing in the R895A mutant, which could possibly be an explanation for high conservation and R895A inactivity. Conversely,
D922 stabilizes the inactive state of OTP86DYW; however, an alanine mutant has reduced activity. We can explain this effect as a result of the inactive ground state (destabilized through D922A) very likely being required for repetitive reactions elicited by a single DYW domain. Mutations of R945, D958 and W960 to alanine show reduced activity, which can be structurally explained by destabilizing effects on OTP86DYW. W960 is tightly embedded in the DYW motif and stabilizes it as it stacks on top of zinc-coordinating H924 beneath highly conserved R918 and maintains a hydrogen bond to the backbone oxygen of V919.

Likewise, D958 stabilizes the DYW motif by formation of a hydrogen bond to highly conserved K928, which explains the impaired function of a respective aspartate mutant in in vivo editing assays with DYW1 and finally our catalytically impaired D958A mutant. S959 in OTP86 (or tyrosine in most DYW domains) points into the solvent, thus, mutation of the corresponding tyrosine to alanine has no effect on DYW1 in vivo activity and the reverse mutation has no effect on OTP86 activity in this work (Fig. 5a); however, a phenylalanine to alanine mutation at this position in Physcomitrium PPR65 showed a severe negative impact on editing. Our structure may help to interpret these past in vivo mutagenesis studies in several ways. Most likely, an impaired stability of the DYW motif as pictured above triggers a destabilization of the active-site Zn1 as they are directly linked via helix α3, which provides residues coordinating Zn1 and Zn2 (Figs. 1b and 2, and Supplementary Fig. 3b,c). In this context, the DYW motif may
also play a role in regulation, for example, the release of the gating domain, repositioning of K915 or binding nucleotides adjacent to the editing site. Finally, we replaced the gating domain’s β-finger residues 875–890 with three glycine residues (Δ875–890GGG). The removal of the entire β-finger markedly reduces the editing activity, which implies its important functional role—probably during

Fig. 5 | Orthogonal in vivo RNA editing validates the OTP86DYW domain structure and activation. a, The C-to-U editing activities at the nad4eU272SL site in E. coli expressed with the PPR56PREL–OTP86DYW fusion protein (OTP86DYW), or its mutants, are plotted. The activities of mutants are relative to that of PPR56PREL–OTP86DYW (82.4 ± 2.1% edited). The bars represent the mean values, with each mutant protein ± s.d. based on three independent experiments (shown as yellow diamonds). The soluble protein expression of each mutated construct in E. coli was verified by western blot analysis (Supplementary Fig. 7). b, The activities of OTP86DYW mutants shown in a plotted on the surface of the inactive OTP86DYW and the activated OTP86DYW* structure as a heatmap (activity is scaled in the bar on the bottom), with untested residues shown in grey.
activation, dimerization or RNA binding—conferred by this region in DYW domains. Although size-exclusion chromatography of isolated OTP86DYW and in the presence of activators (Supplementary Fig. 2a,b) did not indicate dimer formation, prominent protein–protein contacts within the crystal lattice may be physiologically relevant (Supplementary Fig. 8 and Supplementary Note 3)45.

Validation of the OTP86DYW activation mechanism in vitro. We next set out to cross-validate these distinct structural changes of isolated OTP86DYW in solution. The very low amounts of available OTP86DYW led us to conduct differential scanning fluorimetry (DSF). In a typical DSF experiment, an increase of the protein’s melting point ($T_m$) upon ligand binding is observed46. The substrate (CMP), product (UMP) and a K915A mutation do not have an effect on the overall high $T_m$ of OTP86DYW, which is about 71–72 °C in each case (Fig. 6a)47. These results imply a limited accessibility of the active site due to steric inhibition, consistent with the structures and corroborated by mutants. The well-characterized transition-state analogue THU lowers the $T_m$ of OTP86DYW to 60 °C, corresponding to the structural changes we observed following THU co-crystallization. We reason that THU, a potent cytidine deaminase inhibitor, outcompetes the gating domain from the active site, for example, by releasing H890, opening the protein up for substrate access and thereby destabilizing OTP86DYW markedly47,48. Interestingly, this effect is less severe with the K915A mutation, implying a functional role of K915A during activation but not catalysis. Next we asked whether the effect of THU is reversible. A THU-pre-treated OTP86DYW sample was therefore subjected to size-exclusion chromatography to remove THU. Indeed, repurified OTP86DYW resembled the inactive state with a $T_m$ of 72 °C despite THU exposure beforehand. The active state could be restored by the addition of 2 mM THU, resulting in a $T_m$ of 61 °C. We conclude that DYW domains have an inhibited ground state that is restored after the activation and an editing event.

ATP was reported to activate in vitro RNA editing reactions with a recombinant Physcomitrium PPR65 protein as well as

---

**Fig. 6 | In vitro editing and thermal shift assays validate the activation principle of OTP86DYW.** a, Melting points of OTP86DYW and selected variants in the presence of substrate, product and activators. Thermal shift assays were performed in triplicates. Curves were subjected to sigmoidal fitting; error bars indicate the s.d. of the three measurements (shown as yellow diamonds); regular samples are shown as blue bars, whereas SEC-related experiments are shown as orange bars; ‘pre’ indicates a sample pre-incubated with the respective activator and purified by SEC. b, C-to-U conversion for in vitro reactions with recombinant PPR56 or PPR56PPRE1E2–OTP86DYW with the addition of THU, ATP or GTP are displayed in a bar plot. In vitro editing with recombinant PPR56PPRE1E2–OTP86DYW-E894A mutant protein showed no editing activity. Bars indicate the mean value ± s.d. based on three independent experiments (shown as yellow diamonds). A grey dashed line indicates a $T_m$ of 71 °C for wild type protein (WT in a) or the activity of the untreated proteins (in b).
with plant organellar lysates\(^{22,49,50}\); we thus also tested ATP and a concentration of 2 mM was required to drop the \(T_m\) to 65 °C. A very similar picture was obtained by addition of GTP, where 2 mM GTP reduced the \(T_m\) to 65.3 °C. Like THU, the activators ATP and GTP also do not stably bind to OTP86\(^{\text{DYW}}\). When ATP- or GTP-treated OTP86\(^{\text{DYW}}\) is subjected to size-exclusion chromatography, the higher \(T_m\) of the untreated protein (inactive state) is restored in the eluate fractions; however, addition of ATP or GTP to the eluted samples leads again to a decrease of the \(T_m\), indicating a reversible structural change and a stable ground state in the absence of activators (Fig. 6a). In the size-exclusion chromatograms of isolated OTP86\(^{\text{DYW}}\) and OTP86\(^{\text{GTVW}}\) pre-treated with 2 mM ATP, the \(A_{260/280}\) ratios of the respective eluted OTP86\(^{\text{GTVW}}\) peaks are identical (0.54), which further supports the dissociation of the activators from the DYW domain (Supplementary Fig. 2c,d). Contrasting ATP as efficient activator, the addition of 2 mM AMP has a very mild effect. The three phosphates of ATP seem particularly important for activation as the non-hydrolysable analogue AMPPCP had only a mild effect on activation (comparable to AMP); that is, lowering the \(T_m\) to 68 °C. In summary, the activation of OTP86 seems to be either triggered by THU or triphosphate nucleotides.

To gain more insight into whether the H892C mutation effect (activity loss) is of a catalytic or structural nature, we assessed the \(T_m\) of H892C in the presence of the activators. The H892C mutant closely resembles the wild-type protein regarding activation, albeit less pronounced. The detrimental effect of H892C on activity in the in vivo assays therefore relies on the stronger zinc ligand properties of cysteine rather than an impaired structural rearrangement of the catalytic site due to activation (see Fig. 5a).

The L917A mutation showed a prominent decrease in \(T_m\) when THU was added; however, a milder effect with ATP comparable to wild type was observed. This may indicate that the activation via THU and ATP relies on different mechanisms. Finally, the R918A mutant showed a weaker decrease in \(T_m\) in the presence of activator compared with the wild-type. This may be a result of the impaired dimerization capability or an indirect destabilization of the active site via \(\alpha\)-helix 3 and thus reduced activation (see Supplementary Fig. 8).

To consolidate and cross-validate our structural data, in vivo activities and DSF, and to gain more control about the reaction conditions, we conducted in vitro RNA editing assays with purified PPR56 and PPR56\(^{\text{E1E2–OTP86DYW}}\) (Fig. 6b). Contrasting an earlier report, and consistent with the proposed DYW domain activation mechanism, the cytidine deaminase inhibitor THU increases deaminase activity markedly for both proteins in a concentration-dependent manner\(^{22}\). Within this study we were not able to structurally explain this effect due to the absence of THU in the electron density. In agreement with past in vitro editing assays, ATP activates PPR56 and PPR56\(^{\text{E1E2–OTP86DYW}}\) in a concentration-dependent fashion\(^{22}\). We observe that higher ATP concentrations inhibit deaminase activity and thus confirm a highly sensitive regulation of DYW domains by ATP which may be of an allosteric type. Other trinucleotides such as GTP also activate both PPR proteins in a concentration-dependent manner, albeit with a higher sensitivity, confirming the DSF measurements. All assays cross-validate our structural data of the OTP86 DYW domain in its inactive and active states along with a complex regulation mechanism, which suggests an intricate activation of the plant organellar RNA editosome in vivo.

**Discussion**

Our results draw a uniform picture of an unexpected autoinhibition mechanism elicited by DYW domains, which is released in the context of a plant RNA editosome at the site of editing. The data presented here is consistent with past in vivo mutagenesis studies and underlines the cytidine deaminase function of DYW domains in RNA editing\(^{31,32,33,34}\). Typically, cytidine deaminases are highly active enzymes\(^{38}\). With regulated DYW domains, which only exert sensitive regulation of DYW domains by ATP which may be of an allosteric type. Other trinucleotides such as GTP also activate both PPR proteins in a concentration-dependent manner, albeit with a higher sensitivity, confirming the DSF measurements. All assays cross-validate our structural data of the OTP86 DYW domain in its inactive and active states along with a complex regulation mechanism, which suggests an intricate activation of the plant organellar RNA editosome in vivo.

The higher target specificity of DYW type RNA editing factors in plant organelles compared with animal RNA editing deaminase enzymes suggests that the specific binding of RNA by the PPR tract can be a trigger of the DYW activation (Supplementary Note 4). It is also possible that other co-factors in the plant RNA editosome, for example MORF proteins, support moving the gating domain either directly or through changing the conformation of PPR, E1 and E2 domains.

When we compare OTP86\(^{\text{GTVW}}\) to other ligand-bound cytidine deaminases we can extrapolate that the −1 and −2 nucleotide positions relative to the editing site fall into the region of the DYW motif, indicating a head-to-tail arrangement of PPR tract and DYW domain with respect to the direction of the protein sequence of the respective proteins\(^{32,44}\) (Fig. 7). Our observations are in line with a past study in which the 0 to −3 nucleotides bind to the DYW domain, whereas the E1 and E2 motifs do not contribute to binding the target RNA\(^{32}\). In this scenario, the DYW motif bridges the PPR tract and the deaminase/gating domain, which may be the reason for its important structural role within the plant organellar RNA.
Reverse U-to-C RNA editing is observed only in hornwort, most lycophytes and ferns and might be elicited by PPR DYKWP proteins. Our work has several implications that this process may not depend on a strong autoinhibition. We searched for gating domain-like sequences in proteins of all kingdoms using a phmmer search (HmmerWeb version 2.41). Only PPR proteins that included a conserved gating domain sequence were detected. Finally, a comparison of members of the deaminase superfamily identified the gating domain as exclusive insertion in DYW-type PPR proteins.

On the basis of our observations, we propose a regulation mechanism of RNA editing by ATP or other triphosphate nucleotides via the DYW cytidine deaminase activity. RNA editing is directly coupled to the organellar nucleotide metabolism downstream as ATP production is dependent on RNA editing. Conversely, nucleotide levels seem to regulate RNA editing, thus creating a feedback loop. In this scenario, organellar ATP synthesis and RNA editing are mutually regulated to achieve homeostasis. In the light of our artificial in vitro system with isolated proteins, we anticipate a high sensitivity of this feedback loop in vivo possibly owing to the generally low abundance of editing factors observed in mitochondria.

We have further identified a very unusual regulation mechanism involving zinc coordination. In this protein regulation principle, a major domain movement alters the coordination around a zinc atom. In the inactive state, the zinc is inhibited by its coordination setting, which restricts the access of a water molecule as fourth zinc ligand required for catalysis. Upon activation, the DYW gating domain changes its conformation, which triggers the repositioning of a histidine involved in zinc coordination. The altered zinc coordination permits a water molecule to be recruited as a fourth ligand between zinc and the catalytic residue E894 to attack the base for deamination. This regulation principle may also apply to other metalloenzymes beyond DYW deaminases and we are not aware of any similar mechanism described in the current literature.

Our observations explain three decades of previously failed attempts to establish an in vitro RNA editing assay and impaired nucleotide binding of DYW domains. We anticipate our results to be a valuable basis for follow-up experiments for example, a ligand-bound DYW domain structure or cryo electron microscopy studies of a complete editosome. Based on our structure, further in vitro activity assays with structure-guided DYW domain mutants become conceivable where ligand binding, substrate binding and dimerization dependent activation is enhanced or reduced upon mutants.

Methods

Cloning, expression and protein purification. When we set out to determine the structure of a DYW domain, we first screened 18 different A. thaliana (CRR22, CRR28, OTP81, OTP82, OTP84, OTP85, OTP90, LPA66, YS1, RARE1, MEF1, MEF8, MEF10, MEF11, MEF14, MEF22 and MEF29) and three different Physcomitrella patens (PpPPrR_65, PpPPrR_71 and PpPPrR_79) DYW proteins with four different N-terminal starting points (PpPPrR-E12, PpPPrR-E12, PpPPrR-DYW, DYW (according to Cheng et al. and DYW (according to Lurin et al.) and various DYW-containing constructs of nine additional DYW proteins, totalling 113 tested expression constructs. Of these, only one MEF22 and one OTP86 construct yielded small amounts of soluble protein. Only OTP86 (amino acid residues 826–960) crystallized.

A DNA fragment encoding the A. thaliana OTP86 DYW domain (amino acid residues 826–960) was cloned into PET28a to yield a protein (OTP86*), with a tobacco-etch-virus-cleavable (TEV-cleavable) N-terminal Strep-tag. After acid residues 826–960) was cloned into pET28a to yield a protein (OTP86DYW*). The crystals were washed with buffer and fusion proteins were eluted with lysis buffer supplemented with 10 mM dithiobiotin. The eluate was treated with a 1:40 protein mass ratio of TEV protease (in lysis buffer) overnight to remove the N-terminal Strep-tag. Cleaved proteins were further purified via Superdex 75 gel filtration chromatography (GE Healthcare, Unicorn Software 5.20) in 20 mM Tris, pH 7.5 and 150 mM NaCl. Peak fractions of the monomers were pooled, passed over an equilibrated Streptacitin velocity flow column, concentrated to 8–15 mg ml⁻¹, flash frozen in liquid nitrogen and stored at −80 °C. Any alteration to the expression construct described above (for example, variations of the N-terminus length) abolished protein solubility.

Crystallographic analyses. OTP86DYW, supplemented with 2 mM UMP, crystallized by sitting drop vapour diffusion (100 nl protein plus 100 nl reservoir and 30 nl 0.1 M 50% v/v Jefamine M-600 pH 7.0 as an additive) at 4 °C with a reservoir containing 0.1 M glycine, pH 10.5, 1.2 M NaH₂PO₄, 0.8 M K₂HPO₄, and 0.2 M LiSO₄ (space group C2). Crystals were cryoprotected with reservoir solution supplemented with 15% (v/v) ethylene glycol. Diffraction data to 2.5 Å resolution were collected at 100 K at beamline 14.1 of the BESSY II storage ring. All diffraction data were processed with XDS. Activated OTP86DYW*—supplemented with 2 mM CMP and 2 mM THU—crystallized by sitting drop vapour diffusion (1 µl protein plus 1 µl reservoir) in 100 mM sodium acetate (pH 4.6–4.7) and 2 M sodium formate (space group P2₁2₁2₁), with a pronounced degree of translational non-crystallographic symmetry. Crystals were soaked with reservoir solution supplemented with 2 mM CTP, 2 mM THU, and adjusted to a concentration of 3 M sodium formate as a cryoprotectant. Diffraction data to 1.65 Å resolution were collected at 100 K at beamline 14.1 of the BESSY II storage ring. All diffraction data were processed with XDS.

The structure of OTP86DYW was solved by single-wavelength anomalous dispersion with four zinc sites in space group C2 and two molecules per asymmetric unit employing PHENIX.AUTOPOS. The initial density modification map was iteratively improved by manual model building with Coot and refined with PHENIX.REFINE (including experimental phases in the initial stages); automated model building was performed with PHENIX.AUTOBUILD. The structure of OTP86DYW* was solved by molecular replacement with PHASER employing the structural coordinates of a truncated OTP86DYW*, encompassing the deaminase domain and DYW motif. Despite the translational non-crystallographic symmetry, structure solution and refinement were successful, albeit with slightly increased R-factors (see Supplementary Table 1). The remaining model parts were built manually with COOT and with PHENIX.AUTOBUILD in an iterative fashion to improve the model until completion. Structure figures were rendered with open source Pymol v1.8, structural movies were made with Pymol 2.2.3 (Schrödinger) under an academic license. Electrosurface potential was obtained by ABPS employing a Pymol addon. C-alpha r.m.s.d. values for structural comparison were calculated with CCP4i.

DSF. The DSF experiments were performed in a 96-well plate in a plate reader combined with a thermocycler (Stratagene Mx3005P). Purified OTP86DYW or mutants were diluted to 0.2 mg ml⁻¹ buffer A (20 mM Tris, pH 7.5, 150 mM NaCl) supplemented with 100 µM SYPRO orange (1:500 dilution of the stock) in a total volume of 10 µl and pipetted into a 96-well plate. Either 10 µl of buffer A or 10 µl of buffer B supplemented with the respective ligand were added to the SYPRO orange/protein mixture. The temperature was increased from 25 °C to 95 °C and the fluorescence emission was monitored in steps of 1 °C per min with hold steps of 30 s between reads. The fluorescence intensity was then plotted as a function of temperature. The sigmoidal curve from each condition was normalized and corrected for the background signal of the fluorophore in the buffer. The inflection points of the curves, representing the thermal melting temperature of the protein in the respective conditions, were compared. Each experiment was done in triplicate, averaged and a standard deviation of the respective melting temperatures was calculated.

Size-exclusion chromatography. OTP86DYW was analysed by analytical size-exclusion chromatography on a Superdex 75 PC3.2 column (GE Healthcare, Unicorn Software 5.20) in size-exclusion buffer (20 mM Tris, pH 7.5, 150 mM NaCl) at a flow rate of 50–70 µl min⁻¹. Eluted fractions were analysed by sodium dodecyl sulfate–polyacrylamide gel electrophoresis (SDS-PAGE) and the data were converted to DSE Calibration chromatograms for the column were obtained from GE healthcare online support. For clear data presentation, the reverse U-to-C RNA editing was observed only in hornwort, most lycophytes and ferns and might be elicited by PPR DYKWP proteins. Our work has several implications that this process may not depend on a strong autoinhibition.
mercaptoethanol, 0.1% Triton-X-100, 1x complete EDTA-free (Roche) and 1 mM PMSF) and the soluble fraction was isolated after sonication and centrifugation; 7.5µl of the soluble protein lysate was loaded on SDS–PAGE gels for silver staining (Source Data for Supplementary Fig. 7). For the western blot analysis, 150µl of the soluble protein lysate was precipitated with 400 µl of the lysozyme buffer (10 mM Tris-HCl (pH 8.0), 0.1 mM EDTA, 10 mM 1-lysozyme) using a Maxwell RSC Plant RNA Kit system (Promega, www.promega.com). Isolated RNA was used for PCR with reverse transcription (RT–PCR). Data were analysed with Microsoft Excel and plotted with Python/Matplotlib.

Expression of PPR proteins in *E. coli* for in vitro assays. *Escherichia coli* cells from 2 ml culture were resuspended in 1 ml of chilled lysis buffer (50 mM Tris-HCl, pH 7.5, 300 mM NaCl, 10% glycerol, 5 mM imidazole, 0.02% mercaptoethanol, 0.1% Triton-X-100, 1x complete EDTA-free (Roche) and 1 mM PMSF) and the soluble fraction was isolated after sonication and centrifugation; 7.5µl of the soluble protein lysate was loaded on SDS–PAGE gels for silver staining (Source Data for Supplementary Fig. 7). For the western blot analysis, 150µl of the soluble protein lysate was precipitated with 400 µl of aceton. After centrifugation at 4°C for 30 min at 15,000 rpm, the pellet was extracted with 15 µl of 1x loading buffer (50 mM Tris-HCl (pH 6.8), 2% SDS, 0.2% bromophenol blue, 100 mM DTT, 10% glycerol) and loaded onto an SDS–PAGE gel. Expression of recombinant proteins was assayed by western blot analysis with an anti-His-Tag antibody (PGI proteintech Group; AB_11232599) at 1:2,000 dilution followed by incubation with Anti-Mouse IgG, HRP-Linked Whole Ab (GE Healthcare; AB_772209) at 1:50,000 dilution. Signals were detected with ECL Prime Western Blotting Detection Reagent (GE Healthcare) and visualized with an ImageQuant LAS4000 (GE Healthcare). The signal intensities of the western blot analysis were analysed using ImageQuant TL v.8.1 (GE healthcare).

Preparation of RNA editing substrates. Polymerase-chain-reaction fragments were amplified using the pET41K::PPR56 as a template, and primers nad4F/ECrossVGGCGCTCATCACAACATCAATTATTATATAATTCGATAGCTGACCACTCGTCC and nad4RBamH: CGCGCGTAGAGGATCTCAAAATGAGATGATGCTACTATACTATA. This fragment was cloned into pACYC184 digested with EcoRV and BamHI by NEBuilder. Furthermore, using this clone (pACYC184-Pnpnad4) as a template, a PCR amplicon was synthesized with primers T7KS_pACY184EF: GTAAATACGACTCACTATAGGGCTCGAGGTCGACGGTATCAATCTAACAATGGGCTACAT and SKR-pACYC184_EB_R: CCGCTCTAGAATCTAGTGGATCCAGCGGACGGCATTTCTACGAGACGATC and isolated by agarose gel electrophoresis. The PCR amplicon was inserted into pET41K::PPR56 by NEBuilder to generate the intermediate plasmid pET41K::PPR56::nad4. The resulting plasmid was transformed into *E. coli*, and the recombinant protein was extracted and used as the template for in vitro assays.

In vitro RNA editing assay. Standard in vitro RNA editing reaction mixtures contained 100 µM Tris-HCl (pH 7.5), 10 mM maltose, 0.01% mercaptoethanol, 10U of RNaseOUT (Invitrogen), 1 proteinase inhibitor mixture complete EDTA-free (Roche), 100 fmol of mRNA substrate, and 2.5 µg of purified recombinant PPR56 proteins (PPR56nad4* or pET86nad4*) or its mutated variants. The reaction mixtures were incubated at 16°C for 2.5 h and purified RNAs were used for RT–PCR reactions.

Detection of C-to-U RNA editing. Complementary DNA was synthesized with a random hexamer with ReverTra Ace Multiplex with gDNA Remover (Toyobo) for both in *E. coli* and in vitro assays. A reverse primer, upstream of the T7 terminator sequence and a forward primer binding the PPR56 coding region for in *E. coli* assays and KS and SK primers for the in vitro assay were used for RT–PCR amplification with GoTaq Master Mixes (Promega). After 5 min initial denaturation at 94°C followed by 35 cycles each with 30 s denaturation at 94°C, 30 s annealing at 55°C, 1 min synthesis at 72°C. For purification of PCR products, 2U Exol (TAKARA) and 0.5U Shrimp Alkaline Phosphatase (TAKARA) were added and incubated at 37°C for 1 h followed by 15 min at 80°C and sequenced directly (Macrogen, www.macrogen-japan.co.jp or GENEWIZ, https://www.genewiz.com). Sequencing chromatograms were analysed with DNADynano v.1.608 (www.bluetractorsoftware.co.uk). RNA editing was quantified as the ratio of the resulting thymidine peak to the sum of the thymidine and cytidine peak heights at the respective editing site. Editing values are given as the mean of at least three replicates with standard deviations. Data were analysed with Microsoft Excel and plotted with Python/Pyplot.

Reporting summary. Further information on research design is available in the Nature Research Reporting Summary linked to this article.

Data availability
Structure coordinates and diffraction data were deposited with the Protein Data Bank (http://www.pdb.org) under accession codes 7O4E (OTP86nad4*) and 7O4F (OTP86nad4*). Source data are provided with this paper. The data that support the findings of this study are available from the corresponding authors on reasonable request.

Received: 2 November 2020; Accepted: 7 May 2021; Published online: 21 June 2021.

References
1. Hiesel, R., Wissinger, B., Schuster, W. & Brennicke, A. RNA editing in plant mitochondria. *Science* **246**, 1632–1634 (1989).
2. Covello, P. S. & Gray, M. W. RNA editing in plant mitochondria. *Nature* **341**, 662–666 (1989).
3. Gualberto, J. M., Lamattina, L., Bonnard, G., Weil, J. H. & Grienenberger, J. M. RNA editing in wheat mitochondria results in the conservation of protein sequences. *Nature* **341**, 660–662 (1989).
4. Hoch, B., Maier, R. M., Appel, K., Igloi, G. L. & Kössel, H. Editing of a chloroplast mRNA by creation of an initiation codon. *Nature* **353**, 178–180 (1991).
5. Takenaka, M., Zehrmann, A., Verbitskiy, D., Härtel, B. & Brennicke, A. RNA editing in plants and its evolution. *Annu. Rev. Genet.* **47**, 335–352 (2013).
6. Small, I. D., Schallenberg-Rüdinger, M., Takenaka, M., Mireau, H. & Osterter-Biener, B. Plant organellar RNA editing: what 30 years of research has revealed. *Plant J.* **101**, 1040–1056 (2020).
7. Kotera, E., Tasaka, M. & Shikanai, T. A pentatricopeptide repeat protein is essential for RNA editing in chloroplasts. *Nature* **433**, 326–330 (2005).
8. Zehrmann, A., Verbitskiy, D., van der Merwe, J., Brennicke, A. & Takenaka, M. A DYW domain-containing pentatricopeptide repeat protein is required for RNA editing at multiple sites in mitochondria of *Arabidopsis thaliana*. *Plant Cell* **21**, 558–567 (2009).
9. Schmitz-Linneweber, C. & Small, I. Pentatricopeptide repeat proteins: a socket set for organelle gene expression. *Trends Plant Sci.* **13**, 663–670 (2008).
10. Takenaka, M. How complex are the editosomes in plant organelles? *Mol. Plant* **7**, 382–385 (2014).
11. Nakamura, T. & Sugita, M. A conserved DYW domain of the pentatricopeptide repeat protein possesses a novel endoribonuclease activity. *FEBS Lett.* **582**, 4163–4168 (2008).
12. Bousardon, C. et al. Two interacting proteins are necessary for the editing of the NdhH1-1 site in *Arabidopsis thaliana* plastids. *Plant Cell* **24**, 3684–3694 (2012).
13. Yin, P. et al. Structural basis for the modular recognition of single-stranded RNA by PPR proteins. *Nature* **504**, 168–171 (2013).
14. Barkan, A. et al. A combinatorial amino acid code for RNA recognition by pentatricopeptide repeat proteins. *PLoS Genet.* **8**, e1002910 (2012).
15. Andrés-Colás, N. et al. Multiple PPR protein interactions are involved in the recognition of single-stranded RNA by the editosome. *Plant Cell* **24**, 3684–3694 (2012).
16. Barkan, A. & Small, I. Pentatricopeptide repeat proteins in plants. *Annu. Rev. Plant Biol.* **65**, 415–442 (2014).
41. Laitaoja, M., Valjakka, J. & Jänis, J. Zinc coordination spheres in protein.
40. Fica, S. M. & Nagai, K. Cryo-electron microscopy snapshots of the
38. Johansson, E., Neuhard, J., Willemoës, M. & Larsen, S. Structural, kinetic,
37. Diaz, M. F., Bentolila, S., Hayes, M. L., Hanson, M. R. & Mulligan, R. M. A
35. Hayes, M. L., Dang, K. N., Diaz, M. F. & Mulligan, R. M. A conserved
33. Okuda, K., Myouga, F., Motohashi, R., Shinozaki, K. & Shikanai, T.
32. Xiang, S., Short, S. A., Wolfenden, R. & Carter, C. W. The structure of the
43. Xiang, S., Short, S. A., Wolfenden, R. & Carter, C. W. The structure of the
cytidine deaminase-product complex provides evidence for efficient proton
transfer and ground-state destabilization. Biochemistry 36, 4768–4774 (1997).
42. T. A. H. et al. The 1.8 Å resolution crystal structure of the homotetrameric
cytidine deaminase from mouse. Biochemistry 45, 7825–7833 (2006).
45. Krissinel, E. Crystal contacts as nature’s docking solutions. J. Comput.
34. Okuda, K. et al. Pentatricopeptide repeat proteins with the DYW motif have
32. Okuda, K. et al. Quantitative analysis of motifs contributing to the interaction
between PLS-subfamily members and their target RNA sequences in plastid
RNA editing. Plant J. 80, 870–882 (2014).
36. Gerke, P. et al. Towards a plant model for enigmatic U-to-C RNA editing:
the organelle genomes, transcriptomes, editomes and candidate RNA editing
factors in the hornwort Anthoceros agrestis. N. Phytol. 225, 1974–1992 (2020).
46. Niesen, F. H., Berglund, H. & Vedadi, M. The use of differential scanning
fluorimetry to detect ligand interactions that promote protein stability. Nat.
6. 2212–2221 (2007).
37. Costanzi, S. et al. Delineation of the molecular mechanisms of nucleoside
recognition by cytidine deaminase through virtual screening. ChemMedChem
6, 1452–1458 (2011).
38. Cohen, R. M. & Wolfenden, R. Cytidine deaminase from Escherichia coli.
Purification, properties and inhibition by the potential transition state analog
3. 3,4,5,6-tetrahydroxiduridine. J. Biol. Chem. 246, 7561–7565 (1971).
39. Hegeman, C. E., Hayes, M. L. & Hanson, M. R. Substrate and cofactor
requirements for RNA editing of chloroplast transcripts in Arabidopsis
in vitro. Plant J. 42, 124–132 (2005).
40. Takenaka, M. & Bremsicke, A. In vitro RNA editing in pea mitochondria
requires NTP or dNTP, suggesting involvement of an RNA helicase. J. Biol.
6, 47526–47533 (2003).
41. Okuda, K. et al. Pentatricopeptide repeat proteins with the DYW motif have
distinct molecular functions in RNA editing and RNA cleavage in Arabidopsis
chloroplasts. Plant Cell 21, 560–569 (2009).
42. Okuda, K. et al. Quantitative analysis of motifs contributing to the interaction
between PLS-subfamily members and their target RNA sequences in plastid
RNA editing. Plant J. 80, 870–882 (2014).
43. Gerke, P. et al. Towards a plant model for enigmatic U-to-C RNA editing:
the organelle genomes, transcriptomes, editomes and candidate RNA editing
factors in the hornwort Anthoceros agrestis. N. Phytol. 225, 1974–1992 (2020).
44. Kugita, M., Yamamoto, Y., Fujikawa, T., Matsumoto, T. & Yoshinaga, K.
RNA editing in hornwort chloroplasts makes more than half the genes functional.
J. Nut. Acids Res. 31, 2417–2423 (2003).
45. Grewe, F. et al. A unique transcriptome: 1782 positions of RNA editing alter
1406 codon identities in mitochondrial mRNAs of the lycophyte Isoetes engelmannii.
Nat. Acids Res. 39, 2890–2902 (2011).
46. Knie, N., Grewe, F., Fischer, S. & Knoop, V. Reverse U-to-C editing exceeds
C-to-U RNA editing in some ferns—a monilophyte-wide comparison of
chloroplast and mitochondrial RNA editing suggests independent evolution
of the two processes in both organelles. BMC Biol. Evol. 16, 134 (2016).
47. Oldenkott, B., Yamaguchi, K., Tsuji-Tsukinoki, S., Knie, N. & Knoop, V.
Chloroplast RNA editing going extreme: more than 3400 positions of C-to-U
editing in the chloroplast transcriptome of the lycoyte Selaginella uncinata.
RNA 20, 1496–1506 (2014).
48. Glutamates B. et al. The expansion and diversification of pentatricopeptide
repeat RNA-editing factors in plants. Mol. Plant 13, 215–230 (2020).
49. Potter, S. C. et al. HMMER web server: 2018 update. Nat. Acids Res. 46,
W200–W204 (2018).
50. Fukush, P. et al. Single organelle function and organization as estimated from
Arabidopsis. Plant J. 101, 420–441 (2020).
51. Lurin, C. et al. Genome-wide analysis of Arabidopsis pentatricopeptide repeat
proteins reveals their essential role in organelle biogenesis. Plant Cell 16,
2089–2103 (2004).
52. Mueller, U. et al. Facilitates for macromolecular crystallography at the
Helmholtz-Zentrum Berlin. J. Synchrotron Radiat. 19, 442–449 (2012).
53. Kabisch, W. XDS. Acta Crystallogr. D 66, 125–132 (2010).
54. Zwart, P. H. et al. Automated structure solution with the PHENIX Suite.
Methods Mol. Biol. 426, 419–435 (2008).
55. Emsley, P., Lohkamp, B., Scott, W. G. & Cowtan, K. Features and
development of Coot. Acta Crystallogr. D 66, 486–501 (2010).
56. Terwilliger, T. C. et al. Iterative model building, structure refinement and
density modification with the PHENIX AutoBuild wizard. Acta Crystallogr. D
66, 61–69 (2020).
57. McCoy, A. J. et al. Phaser crystallographic software. J. Appl. Crystallogr.
40, 658–674 (2007).
58. Baker, N. A., Sept, D., Joseph, S., Holst, M. J. & McCammon, J. A.
Electrostatics of nanosystems: applications to microtubules and the ribosome.
Proc. Natl Acad. Sci. USA 98, 10037–10041 (2001).
59. Dolinsky, T. J., Nielsen, J. E., McCammon, J. A. & Baker, N. A.
PDB2PQR: an automated pipeline for the setup of Poisson–Boltzmann electrostatics
calculations. Nucleic Acids Res. 32, W665–W667 (2004).
60. Lerner, M. G. & Carlson, H. A. APRS plugin for PyMOL (University of
Michigan, 2006).
61. Potterton, L. et al. CCP4i2: The new graphical user interface to the CCP4
program suite. Acta Crystallogr. D 74, 68–84 (2018).
62. Li, F. W. et al. Anthoceros genomes illuminate the origin of land plants and
the unique biology of hornworts. Nat. Plants 6, 259–272 (2020).
73. Pettersen, E. F. et al. UCSF Chimera—a visualization system for exploratory research and analysis. J. Comput. Chem. 25, 1605–1612 (2004).

74. Barton, G. J. Alscript: a tool to format multiple sequence alignments. Protein Eng. Des. Sel. 6, 37–40 (1993).

75. Kouno, T. et al. Crystal structure of APOBEC3A bound to single-stranded DNA reveals structural basis for cytidine deamination and specificity. Nat. Commun. 8, 1–8 (2017).

Acknowledgements
In remembrance of Axel Brennicke (1953–2017), the co-discoverer of RNA editing in plant organelles and who initiated this project in 2011. We thank B. Gerhardt, D. Pruchner, A. Müller, Y. Yew, N. Holton, O. Ganichkin, K. Vester, A. Bergfort, A. Jörg, M. Burger, J. Wollenhaupt and A. Maeda for excellent technical assistance and experimental support. We are grateful to C. Schmitz-Linneweber, V. Knoop, M. Wahl and T. Shikanai for their valuable support and discussions. We acknowledge access to beamline P13, operated by EMBL Hamburg at the PETRA III storage ring (Hamburg, Germany). We would like to thank I. Bento, G. Boursakov and T. Schneider for their assistance in using the beamline. We acknowledge access to beamlines BL14.1/2/3 of the BESSY II storage ring (Berlin, Germany) via the Joint Berlin MX-Laboratory sponsored by the Helmholtz-Zentrum Berlin für Materialien und Energie, the Freie Universität Berlin, the Humboldt-Universität zu Berlin, the Max-Delbrück Centrum, and the Leibniz-Institut für Molekulare Pharmakologie. This work was supported by DFG grant SCHU 1952/2-2 to M.S.-R., JSPS Grants-in-Aid for Scientific Research no. 18H02462 and DFG grant no. TA624/10-1 to M.T., and start-up funding from the University of Greifswald to G.W.

Author contributions
M.T. and G.W. coordinated and supervised the project. M.T., T.B., S.T., S.H and D.V. cloned OTP86 constructs and performed test expressions. M. S.-R. and B.O. designed initial bacterial expression constructs and developed the bacterial RNA editing assay. S.T. cloned mutants of PPR56P9PE1E2–OTP86DYW, performed the bacterial editing assays and western blot analysis. M.T. and S.T. analysed the data. B.F. expressed, purified and assayed in vitro activity of PPR56P9PE1E2–OTP86DYW. T.B. and G.W. conducted DSF experiments, expressed, purified and crystallized OTP86DYW. G.J.P., C.F. and G.W. collected diffraction data. C.F, M.S.W, G.J.P. and G.W analysed the data. M.S.-R. performed evolutionary conservation analyses. M.T. and G.W wrote the manuscript. All authors discussed the results and commented on the manuscript.

Competing interests
The authors declare no competing interests.

Additional information
Supplementary information The online version contains supplementary material available at https://doi.org/10.1038/s41929-021-00633-x.

Correspondence and requests for materials should be addressed to M.T. or G.W.

Peer review information Nature Catalysis thanks Philippe Giegé, Mamoru Sugita and the other, anonymous, reviewer(s) for their contribution to the peer review of this work.

Reprints and permissions information is available at www.nature.com/reprints.

Publisher’s note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made. The images or other third party material in this article are included in the article’s Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the article’s Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this license, visit http://creativecommons.org/licenses/by/4.0/.

© The Author(s) 2021
Reporting Summary

Nature Research wishes to improve the reproducibility of the work that we publish. This form provides structure for consistency and transparency in reporting. For further information on Nature Research policies, see our Editorial Policies and the Editorial Policy Checklist.

Statistics

For all statistical analyses, confirm that the following items are present in the figure legend, table legend, main text, or Methods section.

- n/a □ Confirmed
- □ The exact sample size (n) for each experimental group/condition, given as a discrete number and unit of measurement
- □ A statement on whether measurements were taken from distinct samples or whether the same sample was measured repeatedly
- □ The statistical test(s) used AND whether they are one- or two-sided
- □ Only common tests should be described solely by name; describe more complex techniques in the Methods section.
- □ □ A description of all covariates tested
- □ □ A description of any assumptions or corrections, such as tests of normality and adjustment for multiple comparisons
- □ □ A description of the full statistical parameters including central tendency (e.g. means) or other basic estimates (e.g. regression coefficient) AND variation (e.g. standard deviation) or associated estimates of uncertainty (e.g. confidence intervals)
- □ □ For null hypothesis testing, the test statistic (e.g. F, t, r) with confidence intervals, effect sizes, degrees of freedom and P value noted Give P values as exact values whenever suitable.
- □ □ For Bayesian analysis, information on the choice of priors and Markov chain Monte Carlo settings
- □ □ For hierarchical and complex designs, identification of the appropriate level for tests and full reporting of outcomes
- □ □ Estimates of effect sizes (e.g. Cohen’s d, Pearson’s r), indicating how they were calculated

Our web collection on statistics for biologists contains articles on many of the points above.

Software and code

Policy information about: availability of computer code

Data collection

Unicorn Software version 5.20 (GE Healthcare) was used to record FPLC data. A Bio-Rad CFX96 Touch Real-Time PCR Detection System was used for DSF data collection. DNAQuant ver. 1.608, https://www.blueactorsoftware.com, was used for sequencing data collection. ImageQuant ver. 8.1, GE healthcare was used for collection of signal intensity from western blot data.

Data analysis

DSF data was analyzed with Bio-Rad CFX Manager 3.1. Microsoft Excel 2016, Python 3.8 / Matplotlib 3.2.2.0 and Kaleidograph 3.5 were used to analyze and display HPLC / DSF data as well as the activities of in vivo and in vitro assays. Structures were displayed with open source PyMol 1.8, structural moves were made with PyMol 2.2.3 (Schrodinger) under an academic license. Academic licences of XDS-Feb. 2001, CCP4 7.0, PHASER 2.3.0 and PHENIX 1.8 with GooT 0.8.92 were used to process and refine the crystallographic data and solve the structure. Microsoft Excel for Mac ver. 16.47.1 was used for calculation for RNA editing efficiency and relative intensity of western blot signals. Structure-based alignment was done with Chimera64 1.14/ Clustal Omega 1.2.2. and shaded with ALCSCRIPT 2.07 d. For phylogenetic analyses, Protein data were extracted from OnekP PPR finder (https://ppr.plantenergy.uwa.edu.au/onekp/) and Phmmer database [https://www.ebi.ac.uk/Tools/hmmer/search/phmmer] aligned with Mega 7.0 [https://www.megasoftware.net/] and MAFFT version 7 [https://mafft.cbrc.jp/alignment/server/]. Conservation plots were generated with Skyline [https://skyline.org/].

For manuscripts utilizing custom algorithms or software that are not central to the research but not yet described in published literature, software must be made available to editors and reviewers. We strongly encourage code deposition in a community repository (e.g. Git-Hub). See the Nature Research guidelines for submitting code & software for further information.
Data

Policy information about availability of data

All manuscripts must include a data availability statement. This statement should provide the following information, where applicable:

- Accession codes, unique identifiers, or web links for publicly available datasets
- A list of figures that have associated raw data
- A description of any restrictions on data availability

Structure coordinates and diffraction data were deposited with the Protein Data Bank (http://www.pdb.org) under accession codes 7OD4 [1P86D]Y2W and 7OD4 [1P86D]Y2W*. The source data underlying figures, 2, 5a, b, 6a, b and Supplementary Figures 1b, c, 2, 3a, and 7 is provided as a Source Data file (data) or as the end of the Supplementary information (uncropped gel pictures and Western Blots). Data employed for structural comparison (PDB entries 1CTU, 1AF2, 1F69, 5KES) were obtained from http://www.pdb.org. Protein Sequence data were obtained from https://www.uniprot.org/[UniProtKB Q0MT1V3, P0C7R1, Q06957, F0D4I5, T21H03, Q9IFS, Q8B8H3, Q0DA86, F56389]. Two Anthoceros angustus Sequence were obtained from [https://www.hornworts.uzh.ch] [Protein ID AagRBON, eev.model.Sc2y5SwM_228.5646.1, AagRBON_eev.model.Sc2y5SwM_368.2386.1]. Protein data used for conservation plot generation were obtained from https://psr.plantenergy.uwa.edu.au/onekp/ and are provided as source data file. Alignments are available as source data file. Reference proteomes available in the phmmer database (https://www.ebi.ac.uk/Tools/pfam/search/phmmer) were screened for gating domain sequences. Accession numbers of output sequences are provided as source data file. Other data that support the findings of this study are available from the corresponding authors upon reasonable request.

Field-specific reporting

Please select the one below that is the best fit for your research. If you are not sure, read the appropriate sections before making your selection.

- Life sciences
- Behavioural & social sciences
- Ecological, evolutionary & environmental sciences

For a reference copy of the document with all sections, see nature.com/documents/nr-reporting-summary-flat.pdf

Life sciences study design

All studies must disclose on these points even when the disclosure is negative.

Sample size

Sample size was not predetermined by any statistical methods, but were chosen according to the standards of the field (at least triplicates). The DSF data analysis, in E.coli and in vitro RNA editing activity assay were performed in triplicates. Western blot analysis was performed once with a standard control of protein amounts by silver staining.

Data exclusions

No data were excluded from the analyses

Replication

All attempts of replication were successful. All experiments were triplicated except for western blot analysis.

Randomization

E.coli colonies used for in vivo and in vitro RNA editing assays were randomly selected and repeated three times. Multiple protein crystals under the same conditions were analyzed.

Blinding

The data were analyzed and checked by multiple authors independently.

Reporting for specific materials, systems and methods

We require information from authors about some types of materials, experimental systems and methods used in many studies. Here, indicate whether each material, system or method list is relevant to your study. If you are not sure if a list item applies to your research, read the appropriate section before selecting a response.

Materials & experimental systems

- n/a
- Antibodies
- Eukaryotic cell lines
- Palaeontology and archaeology
- Animals and other organisms
- Human research participants
- Clinical data
- Dual use research of concern

Methods

- n/a
- Involved in the study
- ChiP-seq
- Flow cytometry
- MRI-based neuroimaging

Antibodies

Antibodies used: 6His, His-Tag Mouse monoclonal antibody, Proteintech. Cat Nr:66005-1-ig, Clone Nr: 1B7G5, AB_2883223 at 1:20,000 dilution. Anti-
| Antibodies used | Mouse IgG, HRP-Linked Whole Ab (GE healthcare) Cat-Nr: NXX931_A8_772209 at 1:50000. |
|----------------|----------------------------------------------------------------------------------|

**Validation**

A specific western signal with a His-tagged protein and no unspecific signal around the size of the recombinant protein in E.coli total proteins were validated in manufacturers’ support website. In addition, we show a control western blot image for E. coli BL21 lysates +/- expressed His-tagged MBP-PPR65 protein in the Source Data for Supplementary Figure 7.