258–269 *Nucleic Acids Research, 2012, Vol. 40, No. 1* *Published online 12 September 2011*
doi:10.1093/nar/gkr740

A conserved lysine residue of plant Whirly proteins is necessary for higher order protein assembly and protection against DNA damage

Laurent Cappadocia, Jean-Sébastien Parent, Éric Zampini, Étienne Lepage, Jurgen Sygusch* and Normand Brisson*

Department of Biochemistry, Université de Montréal, CP 6128, Station Centre-Ville, Montréal H3C 3J7, Québec, Canada

Received March 29, 2011; Revised August 24, 2011; Accepted August 25, 2011

ABSTRACT

All organisms have evolved specialized DNA repair mechanisms in order to protect their genome against detrimental lesions such as DNA double-strand breaks. In plant organelles, these damages are repaired either through recombination or through a microhomology-mediated break-induced replication pathway. Whirly proteins are modulators of this second pathway in both chloroplasts and mitochondria. In this precise pathway, tetrameric Whirly proteins are believed to bind single-stranded DNA and prevent spurious annealing of resected DNA molecules with other regions in the genome. In this study, we add a new layer of complexity to this model by showing through atomic force microscopy that tetramers of the potato Whirly protein WHY2 further assemble into hexamers of tetramers, or 24-mers, upon binding long DNA molecules. This process depends on tetramer–tetramer interactions mediated by K67, a highly conserved residue among plant Whirly proteins. Mutation of this residue abolishes the formation of 24-mers without affecting the protein structure or the binding to short DNA molecules. Importantly, we show that an *Arabidopsis* Whirly protein mutated for this lysine is unable to rescue the sensitivity of a Whirly-less mutant plant to a DNA double-strand break inducing agent.

INTRODUCTION

Plants must protect the integrity of genomic DNA located in three distinct compartments (the nucleus, mitochondria and plastids) against deleterious DNA lesions such as DNA double-strand breaks (DSBs). Although nuclear DNA repair mechanisms have been the subject of intense studies (1,2), very little is known about plant organelle DNA repair pathways (2–4). These pathways must however be robust since both mitochondrion and plastid DNA are routinely exposed to high levels of DNA damaging reactive oxygen species resulting from the intense metabolic activity that takes place in these compartments (5,6).

Until very recently, homologous recombination was the only mechanism proposed to be involved in the repair of DSBs in organelles. Indeed, early studies postulated the existence of a DNA recombination machinery (7,8) based on biochemical evidences such as the presence of functional DNA recombinases in both plastid and mitochondria (9–11), the detection of strand exchange activity in purified organelles (12,13) and the monitoring of recombination events following DNA damage (14,15). DNA repair through recombination is expected to be a very efficient process in both plant mitochondria and plastids as the genomes of these organelles are highly polyploid, thus providing intact DNA molecules that serve as templates for the repair of the damaged DNA (16). Recently, we demonstrated that treatment of *Arabidopsis thaliana* plants with ciprofloxacin, a small molecule that produces DSBs specifically in organelles (17), resulted in the formation of DNA rearrangements in both mitochondria and plastids (15). These DNA rearrangements showed microhomology at their breakpoint junctions, suggesting they are the result of DNA repair through a microhomology-mediated break-induced replication (MMBIR) pathway (15), which is also active in bacteria, yeast and humans (18).

Although most factors involved in the MMBIR pathway have yet to be identified, we recently demonstrated that Whirly proteins protect against MMBIR events in both chloroplasts and mitochondria (15). Whirlies

*To whom correspondence should be addressed. Tel: 514 343 2389; Fax: 514 343 6463; Email: Jurgen.Sygusch@UMontreal.CA
Correspondance may also be addressed to Normand Brisson. Tel: 514 343 2389; Fax: 514 343 6463; E-mail: Normand.Brisson@UMontreal.CA
Present address:
Jean-Sébastien Parent, Institut Jean-Pierre Bourgin, INRA Centre de Versailles-Grignon, 78026 Versailles Cedex, France.

constitute a small family of single-stranded DNA-binding proteins (SSB) that are mainly found in plants (19). They assemble into tetramers resembling whirligigs (20) and perform numerous cellular functions in both nucleus and organelles. Indeed, they were first discovered as transcriptional activators binding the elicitor response element of pathogenesis-related genes in the nucleus of both potato (*Solanum tuberosum*) and *Arabidopsis* (21,22). In *Arabidopsis*, they are also known to bind to telomeres (23) as well as to a distal element upstream of a kinesin gene (24). In maize (*Zea mays*) chloroplasts, WHY1 was shown to bind to both DNA and RNA with a role in intron splicing (25). In barley (*Hordeum vulgare*) chloroplasts, WHY1 is also associated with intron-containing RNA (26). These studies therefore suggest that Whirly proteins might have different functions depending at least in part on intra-cellular localization and developmental stage (15,26).

All plants examined to date possess at least two nuclear genes coding for Whirly proteins. The products of these genes are directed to the mitochondria and plastids, respectively (27). It is not yet known how the Whirly proteins could then move from the organelles to the nucleus (28). In *Arabidopsis*, where there are three genes coding for Whirly proteins, WHY1 and WHY3 are directed to the chloroplasts whereas WHY2 is directed to the mitochondria. In absence of both WHY1 and WHY3, DNA rearrangements mediated by microhomologies occur spontaneously in the chloroplasts and often correlate to the development of white–yellow variegated leaves (29). Induction of DSBs by treatment with the gyrase inhibitor ciprofloxacin leads to a further increase in microhomology-mediated DNA rearrangements in the chloroplasts of *why1 why3* mutant plants that lack both *WHY1* and *WHY3* genes (15). Contrasting with the situation in *why1 why3* plants, the inactivation of *WHY2* does not result in any visible or molecular phenotype in absence of DNA stress (30). Upon induction of DSBs, however, plants lacking WHY2 accumulate microhomology-mediated mitochondrial DNA rearrangements more readily than wild-type *Arabidopsis* plants (15). The different behavior between *why2* and *why1 why3* plants could be due to the partial functional redundancy provided by other SSBs present in plant mitochondria (15,31). Whirly proteins are believed to modulate the MMBIR pathway either by binding the resected DNA that forms during processing of DSBs, thus favoring error-free repair through homologous recombination (15), or by binding ssDNA from all over the genome (25,29,30) and thus preventing spurious annealing of DNA through microhomologous sequences (15).

In contrast to most SSBs that use an OB fold as their main ssDNA-binding platform (32), Whirly proteins use a Whirly domain (19,20). This domain is typically composed of two four-stranded β and two α helices. These helices constitute the core of the protein against which stacks the β-sheets in a whirligig-like manner (20). Determination of the atomic structure of the potato Whirly protein WHY2 in complex with different ssDNA molecules shed light on the mechanism of DNA binding (15). WHY2 binds ssDNA by establishing hydrogen bonds with the phosphate moiety of the ssDNA and by stacking nucleobases in between the β-sheets of adjacent subunits. This enables binding of ssDNA with high affinity and little sequence specificity as the bases are not brought in close contact with the protein surface. SSBs can bind DNA with various degrees of cooperativity (33,34). For bacterial SSB, the cooperative binding to long ssDNA molecules is due to the formation of a protein filament (34,35). In the case of Whirly proteins the formation of a superstructure upon binding long DNA molecules has not yet been investigated.

A KGKAAL sequence motif is a hallmark of Whirly proteins as it is highly conserved among plant Whirlies (19,20). This motif is important for ssDNA binding as its replacement with an alternate sequence abolishes the interaction with ssDNA (20). Rather than interacting directly with DNA, many residues within this motif seem important to maintain the stability of the β-sheet that contains numerous DNA-interacting residues. The role of the second lysine of the KGKAAL motif is however enigmatic. Indeed, in the crystal structure of the potato WHY1 protein (PDB 1L3A), the ε amino group of this conserved residue points in the general direction of the solvent in three out of four subunits and does not appear to contribute to the overall stability of the tetramer.

In this article, we show that WHY2, a *Solanum tuberosum* Whirly protein located in the mitochondria, binds long ssDNA in a cooperative manner through hexamerization of Whirly tetramers. This results in the formation of hollow protein shells of 12 nm diameter with 432 symmetry. We also highlight the unique contribution of the second lysine residue of the KGKAAL motif to the formation of 24-mers. Mutation of the equivalent lysine in At-WHY1 partially abolishes the complementation of the etiolated/variegated phenotype in *why1 why3 Arabidopsis* plants treated with ciprofloxacin, highlighting the functional importance of this residue *in planta*.

MATERIALS AND METHODS

Preparation of WHY2 and WHY2 K67A

Details concerning cloning, expression and purification of WHY2 are reported elsewhere (36). The K67A mutant of WHY2 was generated using the method developed by Zheng (37). The mutation was confirmed by DNA sequencing. Expression and purification were performed as for the non-mutated protein.

Data collection and structure determination

WHY2 K67A in the free form (forms I and II; 20 mg/ml) was crystallized by the hanging drop vapor diffusion method using a precipitant solution containing 100 mM MOPS pH 7.0, 18–24 % (*w*/*v*) PEG1000 and 100 mM $NH_4H_2PO_4$. The WHY2 K67A-dT_{32} (10 mg/ml) complex was crystallized by the hanging drop vapor diffusion method using a precipitant solution containing 100 mM Tris–HCl pH 8.0, 15 % (w/v) PEG6000 and 1.2 M LiCl. Diffraction data were collected using an ADSC Quantum 315 detector at beamline X29 at the National Synchrotron

Light Source (NSLS) at the Brookhaven National Laboratory (BNL, USA) or using a Rayonix MX300 detector at beamline 08ID-1 at the Canadian Light Source (CLS). Initial phases for WHY2 in the free form and the WHY2 K67A–dT_{32} complex were obtained by molecular replacement using WHY2 (PDB 3N1H) and WHY2-dT_{32} (PDB 3N1I) as search templates. Phases were improved by iterative cycles of model building with Coot (38) and refinement with CNS (39,40) and Phenix (41). Clear density was visible for protein residues 55–215 except for WHY2 K67A in the free form II. In this structure, clear density was visible for residues 54–141 and for residues 150–214. The residues 142–149 were omitted as they are disordered. Test data sets were randomly selected from the observed reflections prior to refinement. The figures were prepared with PyMOL (42).

Electrophoretic mobility shift assays

The apparent dissociation constant for WHY2 K67A proteins bound to DNA was determined by EMSA as described previously (15). The autoradiograms were scanned and the intensity of the bands corresponding to bound and free radiolabeled ssDNA was quantified using ImageJ (NIH, http://rsb.info.nih.gov/ij/). The data was plotted and analyzed using Prism 5 Demo (GraphPad Software). The apparent dissociation constants were calculated by fitting the plot of the fraction of ssDNA bound versus protein concentration to the Hill equation.

Agarose-based electrophoretic mobility shift assays

Binding reactions including 600 ng of M13mp18 ssDNA (USB) and the serially diluted WHY2 or WHY2 K67A protein were incubated for 20 min at 20°C in a buffer containing 20 mM Tris–HCl pH 8.0 and 20 mM NaCl. Reactions were run on a 0.7 % (w/v) agarose gel with 40 mM Tris, 20 mM acetic acid, 1 mM EDTA pH 8.0 and 0.005 % (v/v) ethidium bromide at 20°C under 80 V constant voltage for 1 h. DNA and protein–DNA complexes were visualized by UV transillumination.

Atomic force microscopy

Atomic force microscopy (AFM) was performed in tapping mode in air using a JEOL JSPM-5200 microscope. Topographic and phase images of 512 × 512 pixels with a scan size of 1.5 × 1.5 μm were simultaneously acquired at a scan frequency of 1–1.5 Hz. Topographic images were used for height-measurement whereas phase images were used to identify features on the sample surface that were not easily distinguished by height (43). Raw AFM images were processed using a JEOL software package (WinSPM DPS software, JEOL Ltd.) and the Gwyddion software (Czech Metrology Institute, http://gwyddion.net/).

Imaging of the protein–DNA complexes was performed as described (44) using a 5-μl droplet of a solution containing 1 μg ml^{-1} WHY2 or WHY2 K67A, 20 mM Tris–HCl pH 8.0, 20 mM NaCl and 50 μM spermidine. Imaging of the free protein was performed using a 5-μl droplet of a solution containing 1 μg ml^{-1} WHY2, 20 mM Tris–HCl pH 8.0, 20 mM NaCl and 100 μM $NiCl_2$. All samples were deposited onto the surface of freshly cleaved mica for 1 min. The surface was then rinsed with a 0.02% uranyl acetate solution. The samples were then rapidly rinsed with pure water (Millipore) to obtain a clean surface after drying under a stream of nitrogen.

Plant material and growth conditions

The *why1 why3* mutant line was described previously (15,29). To complement *why1 why3*, the entire coding sequence of WHY1, or its K91A mutant, was amplified by RT-PCR using total RNA from *A. thaliana* ecotype Col-0 and cloned into the pGreen vector containing the BAR casette and under the control of the 35S CaMV promoter (45). These constructs were individually co-transformed with the pSOUP vector into a GV3101 pMP90 *Agrobacterium tumefaciens* strain. *Arabidopsis why1 why3* plants were then transformed by the floral dip method as described (46). Seeds were harvested, planted on soil and screened for BASTA resistance. Two independent homozygous lines were obtained for each construct. The *why1 why3* #1 and the *why1 why3* #2 lines were derived from *why1 why3* 35S:WHY1 #1 and *why1 why3* 35S:WHY1 #2 following segregation of the transgene. For the complementation analysis, sterilized seeds of Col-0 (wild type), *why1 why3* and complemented lines were grown for 10 days on Murashige and Skoog basal medium (Sigma-Aldrich) containing 0.125 μM ciprofloxacin. One-way ANOVA and Tukey's HSD test were used for multiple phenotype comparisons. Plants were analyzed by protein gel blot and eletrophoretic mobility shift assay as described previously (29).

RESULTS

WHY2 assembles into hexamers of tetramers in the crystals

Analysis by the Protein Interfaces, Surfaces and Assemblies (PISA) service (47) indicates that, in all the WHY2 crystal structures, WHY2 tetramers further assemble into stable hexamers of tetramers, or 24-mers. K67, the second lysine of the KGKAAL sequence motif, contributes to the formation of this assembly as it forms a hydrogen bond with the backbone carbonyl of the F138 residue of a symmetry-related tetramer in the crystal structure of the potato WHY2 protein (PDB 3N1H) or a hydrogen bond with the edge of a nucleobase of a symmetry-related tetramer in the crystal structures of three different WHY2–DNA complexes (PDB 3N1I, 3N1J, 3N1K, 3N1L).

The 24-mers form roughly spherical protein shells with 432 symmetry (Figure 1A and 1B) that measure 12 nm of diameter, have a 5 nm diameter hollow core, and have both their N- and C-termini projecting outward from the sphere. Each protein shell has six 0.8 nm-wide pores that coincide with the 4-fold axes. These pores allow access to an electropositive inner cavity (Supplementary Figure S1). Adjacent 24-mers interact with each other through their β-strands. In DNA-bound WHY2 complexes, binding of ssDNA takes place on the external surface of the spheres (Figure 1C).

We previously hypothesized that long ssDNA molecules could travel from one binding site to another within a

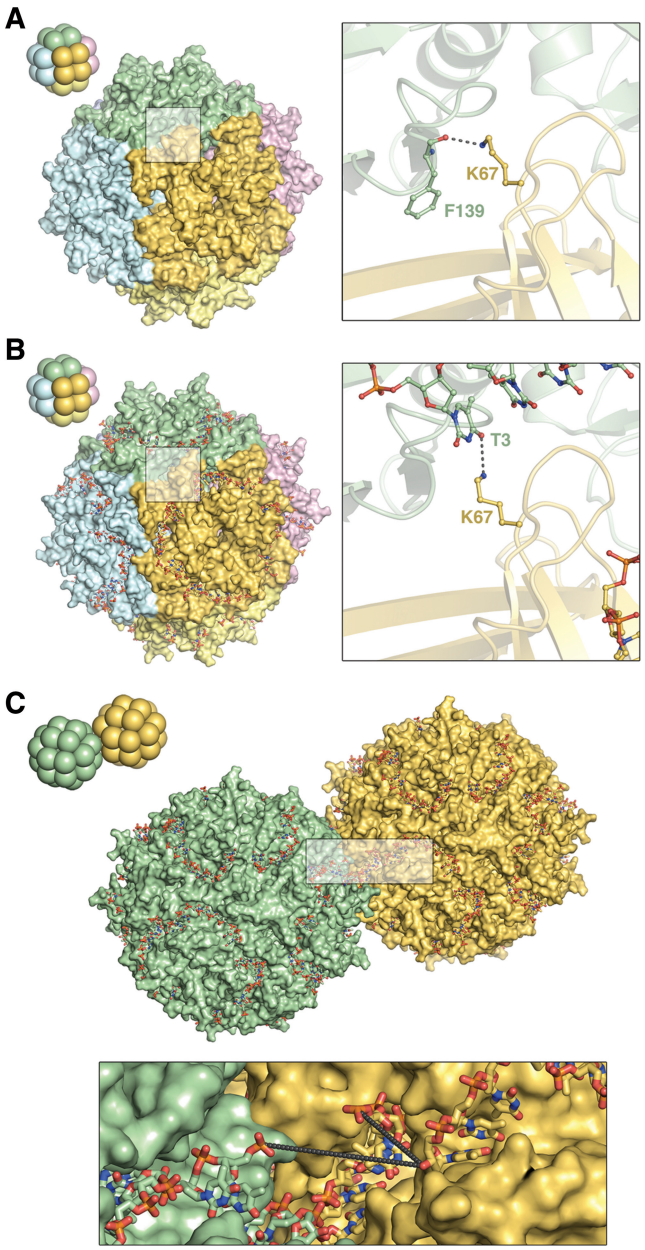

Figure 1. Critical role of K67 in the assembly of tetramers into 24-mers. (**A**) Left, surface representation of the 24-mer assembly, obtained by applying crystallographic symmetry along the 4- 3- and 2-fold axes, in the free form structure of WHY2 (PDB 3N1H). Individual tetramers are colored in different colors. Right, interaction between K67 and the backbone of F138 in the same structure. Protein residues are in stick representation. Hydrogen bonds are represented as black dashes. Top left, schematic of the 24 subunits drawn as spheres. (**B**) Left, surface representation of the 24-mer assembly in the WHY2-ERE$_{32}$ complex structure (PDB 3N1I). The nucleotides are in stick representation. Right, interaction between K67 and T3 in the same structure. The presentation and orientation is similar as in (A). (**C**) Top, overall view of two adjacent 24-mers in the crystal of WHY2-ERE$_{32}$. The hexamers, depicted in surface representation are colored in green and yellow. Bottom, interface between two 24-mers showing the close vicinity of DNA fragments. DNA molecules are in stick representation. The black dash lines illustrate the two possible routes for the DNA. Top left, schematic of the two 24-mers with individual subunits drawn as spheres.

single tetramer (15). Indeed, although WHY2 was crystallized with a 32 nucleotide long ssDNA molecules, each protein protomer only binds a nine nucleotide long fragment (15). Interestingly, the 5′- and 3′-ends of DNA fragments bound by adjacent 24-mers are separated by only 1.9 nm (Figure 1C). Three nucleotides can easily span this distance considering an average of 0.6 nm between phosphorus atoms of neighboring nucleotides. Our present observation that DNA ends are located in close proximity between adjacent 24-mers thus raises the possibility that two pathways exist for the ssDNA in the crystals: one that goes to the adjacent protomer of the same tetramer and one that goes to an adjacent 24-mer (Figure 1C).

WHY2 assembles into hexamers of tetramers *in vitro* upon binding long DNA

To verify whether WHY2 assembles into 24-mers *in vitro*, we observed unbound and dT$_{32}$-bound WHY2 proteins using AFM, a technique that allows estimation of the oligomeric state of particles through measurement of their heights. For WHY2 in the free form or in complex with a dT$_{32}$ oligonucleotide, only scattered particles with an average height of 3–5 nm were detected by AFM (Figure 2A and 2B). These particles are likely tetramers lying flat on the mica surface, consistent with a height of 4 nm for the WHY2 tetramer in the crystals and with the results of size-exclusion chromatography showing that WHY2 and WHY2-dT$_{32}$ complexes mainly exist as tetramers (Supplementary Figure S2). This indicates that, *in vitro*, WHY2 does not form 24-mers constitutively or upon binding small DNA molecules. In contrast, AFM revealed that complexes of WHY2 bound to M13mp18, a 7249 nucleotide long ssDNA molecule, existed as grouped particles (Figure 2C and 2D). These particles exhibited a bimodal distribution with average heights of 6.0 and 10.4 nm. A height of 6.0 ± 1.7 nm is similar to the heights measured for WHY2 (5.1 ± 2.0 nm) and WHY2-dT$_{32}$ (3.4 ± 1.4 nm) suggesting that these species could be DNA-bound tetramers while a height of 10.4 ± 1.9 nm is in agreement with that of the 24-mers observed in the crystal lattice of WHY2 (12 nm). This suggests that 24-mers of WHY2 may form *in vitro* either at very high local concentration (the concentration of Whirly proteins reaches 35 mM in the crystals) or upon binding long single-stranded DNA molecules.

K67 stabilizes the 24-mer assembly of WHY2

To assess the contribution of K67 to the stability of the 24-mers, we produced a K67A mutant of WHY2. This mutant assembles into tetramers both in the free form and upon binding of a dT$_{32}$ oligonucleotide (Supplementary Figure S2). Electrophoretic mobility shift assays (EMSA) experiments further reveal that this mutant binds dT$_{32}$ with an apparent dissociation constant of 3.6 ± 0.5 nM (Figure 3A). This is similar to the value observed for the binding of the non-mutated protein to dT$_{32}$ [4.5 ± 0.9 nM; (15)], thus suggesting that the K67A mutation does not affect the binding of individual tetramers to short ssDNA molecules. The binding of

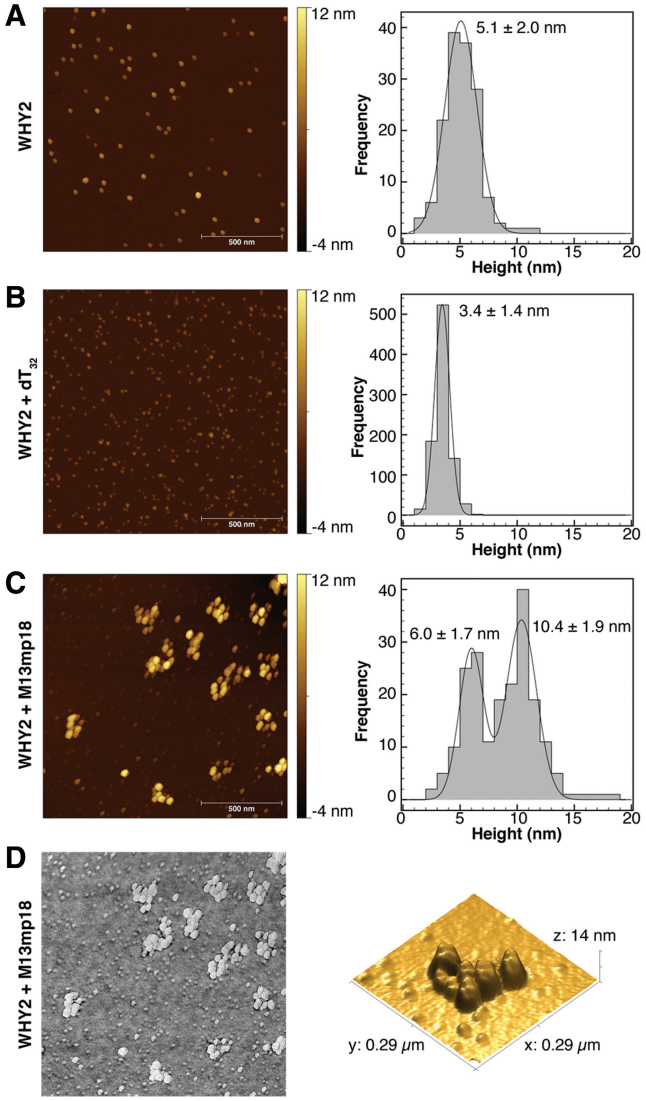

Figure 2. Whirly proteins form 24-mers *in vitro* upon binding long ssDNA molecules. (**A**) Topographic imaging of WHY2 in the free form on a mica surface obtained by AFM (left) and height measurement of the particles (right). Particle heights of 5.1 ± 2.0 nm and 5.4 ± 1.2 nm were measured in two independent experiments. (**B**) Topographic imaging of WHY2–dT_{32} complexes on a mica surface obtained by AFM (left) and height measurement of the particles (right). Particle heights of 3.4 ± 1.4 nm and 3.2 ± 1.6 nm were measured in two independent experiments. (**C**) Topographic imaging of WHY2–M13mp18 complexes on a mica surface obtained by AFM (left) and height measurement of the particles (right). Particle heights of 6.0 ± 1.7 nm and 5.6 ± 1.8 nm (first population) and 10.4 ± 1.8 nm and 8.9 ± 0.7 nm (second population) were measured in two independent experiments. (**D**) Phase imaging of WHY2–M13mp18 complexes on a mica surface obtained by AFM (left) and zoom on WHY2–M13mp18 complexes (right). All acquisitions were performed in tapping mode in air.

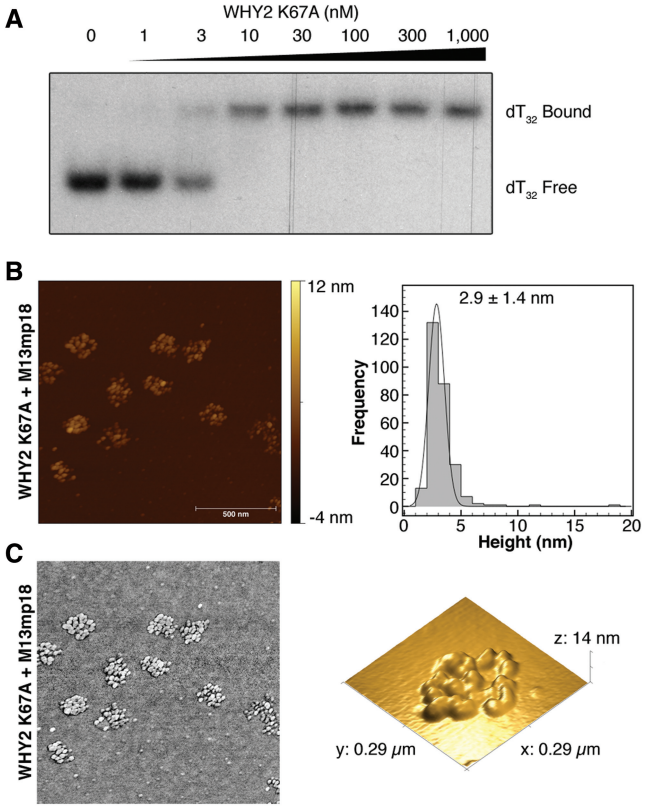

Figure 3. A WHY2 K67A mutant does not assemble into 24-mers *in vitro*. (**A**) Representative EMSA results showing the binding of WHY2 K67A to $(dT)_{32}$. Increasing amounts of WHY2 were incubated with the target oligonucleotide and the complexes resolved on a 10% (w/v) polyacrylamide gel. (**B**) Topographic imaging of WHY2 K67A–M13mp18 complexes on a mica surface obtained by AFM (left) and height measurement of the particles (right). Particle heights of 2.9 ± 1.4 nm and 2.6 ± 0.8 nm were measured in two independent experiments. (**C**) Phase imaging of WHY2–M13mp18 complexes on a mica surface obtained by AFM (left) and zoom on WHY2–M13mp18 complexes (right). All acquisitions were performed in tapping mode in air.

WHY2 K67A to M13mp18 ssDNA was then assessed by AFM (Figure 3B and C). As for WHY2-M13mp18, the complexes existed as grouped particles. However, the average height of the particles of the WHY2 K67A–M13mp18 complex (2.9 ± 1.4 nm) was much lower than for WHY2–M13mp18 (10.4 ± 1.9 nm) and was consistent with that of DNA-bound tetramers lying flat on the mica surface (predicted size of 4 nm). This suggests that WHY2 K67A does not assemble into 24-mers upon binding long DNA.

To ensure that the K67A mutation does not affect the overall structure of the protein, we solved the crystal structure of WHY2 K67A both in the free form (Figure 4A and Table 1) and in complex with a dT_{32} oligonucleotide (Figure 4B and Table 1). WHY2 K67A-dT_{32} only crystallized in the F432 space group whereas WHY2 K67A crystallized in two different space groups under the same chemical conditions: F432 (form I) and $P42_12$ (form II). The proteins in both space groups are structurally similar to the non-mutated protein (Figure 4A and 4B), thereby confirming that the K67A mutation does not perturb the overall structure of the protein. A small rearrangement can however be observed in the loop consisting of residues 139–145 in the free form structure of WHY2 K67A (form I) as compared to the non-mutated protein (Figure 4A). In the absence of K67, this loop

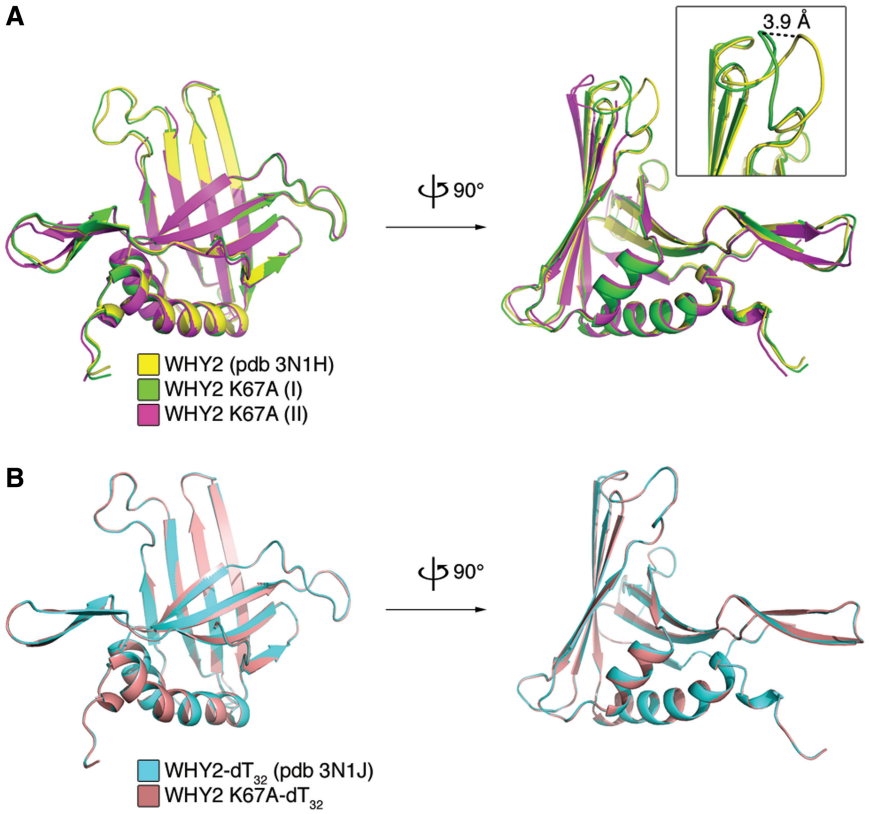

Figure 4. Structural comparison of the K67A variant with the non-mutated WHY2 protein. (**A**) Comparison of the two crystal forms of WHY2 K67A with the non-mutated protein. The RMSD for superimposing the crystal forms I and II of WHY2 K67A onto the non-mutated protein are 0.7Å and 1.1Å, respectively. (**B**) Comparison of the dT_{32}-bound form of WHY2 K67A with the dT_{32}-bound non-mutated protein. The RMSD for superimposing WHY2 K67A onto the non-mutated protein is 0.2Å.

adopts a new conformation that enables it to interact with a symmetry-related tetramer. Interestingly, WHY2 K67A tetramers are arranged in 24-mers in both WHY2 K67A (form I) and WHY2 K67A-dT_{32} crystal lattices suggesting that the K67A mutation does not preclude the capacity of WHY2 K67A tetramers to assemble into 24-mers. Consistent with the EMSA results, the dT_{32}-bound K67A mutant adopts the same conformation as the non-mutated protein (Figure 4B) with a root-mean square deviation (RMSD) of merely 0.2 Å for all aligned Cα, explaining why the binding to dT_{32} is not affected by this mutation.

Cooperative binding of WHY2 upon binding long DNA molecules

The close proximity of the DNA ends in the crystal structure of WHY2–DNA complexes, along with the existence of interactions between adjacent 24-mers, suggest that WHY2 could bind ssDNA in a cooperative manner. To verify this hypothesis, we performed an agarose-based EMSA experiment that enable qualitative evaluation of the cooperative binding of proteins to long DNA molecules (48). Increasing amounts of WHY2 were incubated with a fixed amount of M13mp18 ssDNA and the migration profile was analyzed on an agarose gel (Figure 5A). As the concentration of WHY2 was increased, DNA became trapped in a large protein–DNA complex that barely entered the gel (Figure 5A). This suggests that WHY2–ssDNA complexes form in an all-or-nothing manner characteristic of cooperative binding (48–50). In contrast, agarose-based EMSA experiments show that the K67A mutation abrogates the cooperative binding of WHY2 to ssDNA, as complexes of decreasing mobility were formed when increasing amounts of WHY2 K67A were added to M13mp18 (Figure 5B). Globally, these results indicate that the K67A mutation can disrupt both the 24-mer assembly and the cooperative binding to long ssDNA.

Analysis of the K91A mutation in *Arabidopsis*

Several studies investigated the inclusion of Whirly proteins as part of DNA- or RNA-containing complexes (25,26,29,51). Similarly to the results obtained in maize (25), size exclusion chromatography reveals that, in extracts enriched in *Arabidopsis* chloroplasts, part of the WHY1/3 population is found in large complexes that contain DNA (Supplementary Figure S3). We then sought to test the contribution of the lysine residue to the DNA-repair function of the Whirly proteins *in vivo*. Since the KGKAAL motif is conserved throughout the Whirly protein family, we took advantage of the *why1 why3* mutant plants that we previously isolated (29).

Table 1. Data collection and refinement statistics

Protein	WHY2 K67A Free Form I	WHY2 K67A Free Form II	WHY2 K67A–dT_{32}
Beamline/wavelength	CLS-08ID-1/0.98 Å	NSLS-X29/1.29 Å	NSLS-X29/1.08 Å
Space group	F432	$P42_12$	F432
Cell dimensions *a*, *b*, *c* (Å)	168.37, 168.37, 168.37	90.58, 90.58, 50.93	166.47, 166.47, 166.47
Resolution (Å)	50–2.35 (2.48–2.35)	50–1.78 (1.84–1.78)	50–2.45 (2.54–2.45)
Total reflections	201 296	258 074	135 360
Unique reflections	9012	20443	7685
R_{sym} (%)	5.9 (102.5)	4.8 (31)	6.7 (67)
$I/\sigma I$	31.0 (2.8)	21.5 (2.6)	24.5 (2.3)
Completeness (%)	100.0 (100.0)	99.4 (95.3)	99.6 (97.8)
Multiplicity	22.3 (18.6)	12.6 (5.8)	17.6 (7.0)
Refinement statistics			
Resolution (Å)	50–2.35	50–1.78	50–2.45
Reflections (total/test)	9012/541	19650/960	7469/586
R_{work}/R_{free} (%)	22.36/26.64	19.21/21.57	24.98/26.73
Number of atoms			
Protein	1263	1220	1263
DNA/Phosphate	10	10	180
Water	42	209	39
B-factors			
Protein	60.8	37.7	68.1
DNA/Phosphate	69.0	49.0	75.4
Water	60.1	49.8	66.2
RMSDs			
Bond lengths (Å)	0.005	0.007	0.003
Bond angles (°)	0.894	1.000	0.705
Ramachandran[a]			
Favored (%)	98.1	98.7	94.3
Outliers (%)	0	0	0

Values in parentheses are for highest-resolution shell.
[a]MolProbity analysis (73).

Indeed, the absence of Whirly proteins in the plastids of these plants makes them sensitive to low doses of ciprofloxacin, as they develop extensive variegation/etiolation whereas WT control is only marginally affected (15). We therefore complemented the absence of whirlies by introducing either At-WHY1 or At-WHY1 K91A (the equivalent of K67A in St-WHY2) transgene into the mutant background. Consistent with our previous results, upon treatment with 0.125 µM ciprofloxacin, only 1% of WT plants developed variegated/etiolated leaves whereas, in the same conditions, 52 and 59% of two *why1 why3* lines tested developed variegation/etiolation (Figure 6A). Interestingly, two *why1 why3* lines complemented with *35S:WHY1* displayed partial rescue of the Whirly function with 18 and 10% of leaves showing variegation/etiolation. In contrast, the leaves of two *why1 why3* lines genetically transformed with *35S:WHY1 K91A* displayed 38 and 43% variegation/etiolation, respectively, which is more similar to the levels observed for the untransformed *why1 why3* lines (Figure 6A). Noteworthy, there are no significant differences ($P > 0.05$) between the two independent lines obtained for each mutant (see Supplementary Table S1 for a complete statistical treatment of these data). There is however a significant difference ($P < 0.001$) between the variegation/etiolation percentage observed for *35S:WHY1* and the *35S:WHY1 K91A* line thus underlining the critical role of K91 *in vivo*. The same tendencies were observed in a second independent experiment (Supplementary Figure S4). Figure 6B demonstrates that the WHY1 protein accumulates to similar levels in *35S:WHY1* and *35S:WHY1* K91A lines, far above the endogenous level. Figure 6C indicates that the ssDNA-binding activity is also similar in *35S:WHY1* and *35S:WHY1* K91A lines and proportional to the protein level. Thus, consistent with our *in vitro* analyses of St-WHY2, no reduction in DNA-binding activity was observed upon K91A mutation. Globally, these results suggest that the K91A mutation does not destabilize or reduce the DNA-binding affinity of WHY1. Nevertheless, this mutation interferes with the biological function of Whirly proteins in DNA repair.

DISCUSSION

We previously demonstrated that Whirly proteins protect *Arabidopsis* organelle DNA against the detrimental effects of ciprofloxacin by preventing DSB repair by the error-prone MMBIR mechanism (15). We now show that this function is compromised when the second lysine residue of the KGKAAL motif of At-WHY1 is mutated. Indeed, when we introduce an At-WHY1 transgene into the ciprofloxacin-sensitive *why1 why3* background, we are able to rescue, for the most part, the plant resistance to the antibiotic. The fact that we are not able to fully rescue the resistance to ciprofloxacin might be due to a difference in the timing of expression or in spatial localization of the protein as a consequence of the strong viral promoter that was used. It remains, however, that when we express the

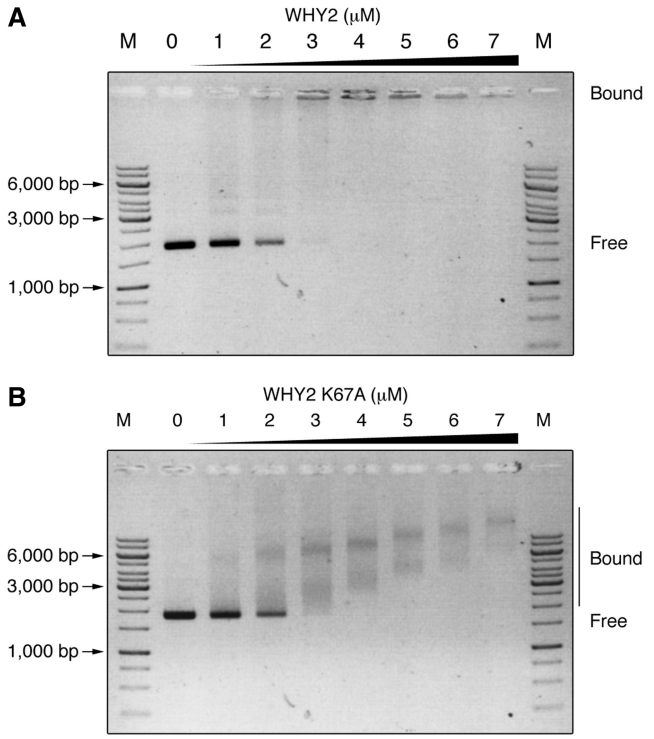

Figure 5. Cooperative binding of WHY2 upon binding long ssDNA molecules. (**A**) Black/white inverted image of an agarose-based EMSA showing cooperative binding of WHY2 to M13mp18. (**B**) Black/white inverted image of an agarose-based EMSA showing non-cooperative binding of WHY2 K67A to M13mp18. The DNA was incubated with increasing amount of WHY2 or WHY2 K67A and the complexes were resolved on a 0.7% (w/v) agarose gel containing ethidium bromide. DNA and protein–DNA complexes were visualized by UV transillumination. M represents the molecular weight markers. 'Bound' indicates the position of the protein–DNA complexes that barely entered the gel. 'Free' indicates the unbound DNA.

K91A mutated transgene into the mutant plants, we fail to rescue the resistance to the antibiotic. This is at first glance intriguing since this mutation does not affect WHY1 stability nor its binding affinity to short ssDNA oligonucleotides. However, the structural analysis of St-WHY2, an At-WHY1 paralogue, provides clues for understanding the possible function of this residue. Indeed, we established that the Whirly proteins are arranged in 24-mers in the crystals of WHY2 or WHY2–DNA complexes and that the second lysine residue of the KGKAAL motif contributes to the stabilization of this assembly. Furthermore, we confirmed that the 24-mer assembly, which forms *in vitro* upon binding long DNA sequences, is abolished when the K67 residue is mutated. Therefore, a possible role for this residue *in planta* could be to promote 24-mer assembly and cooperative binding to DNA. These could be general properties of plant Whirly proteins as both the Whirly fold and the second lysine residue of the KGKAAL motif are strongly conserved among plant Whirlies (19,20).

AFM was used on several instances to monitor a change in the oligomeric state of proteins upon DNA binding. For example, AFM revealed that APOBEC3G, a human ssDNA cytosine deaminase that mainly exists as a monomer in solution, forms dimers upon ssDNA binding (52). Whereas certain nucleoid-associated proteins maintain DNA in an accessible state by counter-acting DNA aggregation (53), other proteins tend to aggregate upon DNA binding (54). AFM indeed revealed that CbpA, a bacterial nucleoid-associated DNA-binding protein which in free form exists in a monomer–dimer equilibrium, forms large aggregates upon binding long double-stranded DNA molecules (55). Likewise, DPS (DNA-protecting Protein during Starvation), another bacterial nucleoid-associated DNA-binding protein that constitutively forms dodecamers, also produces large aggregates upon binding long double-stranded DNA molecules (56). The present study reveals that upon ssDNA binding, tetramers of WHY2 form superstructures whose size is consistent with that of the 24-mers observed in the crystal lattice. However, contrarily to other DNA-binding proteins monitored by AFM, WHY2 does not form large aggregates upon DNA binding. Still, the packed-beads appearance of WHY2–DNA complexes suggests that WHY2 binding induces DNA condensation. This packing could be mediated by protein–protein interactions in between adjacent 24-mers.

Similarly to plant Whirlies, bacterial SSB proteins bind ssDNA with high positive cooperativity through the formation of a higher-order protein assembly (35,57). In *Escherichia coli*, the SSB protein can exist in two states, $(SSB)_{35}$ and $(SSB)_{65}$, that bind ssDNA with different degrees of cooperativity (58). Studies have revealed that the high positive cooperativity state $(SSB)_{35}$ is stabilized by the W54S mutation (59,60). Interestingly, when the gene coding for this mutated protein is expressed in *E. coli*, cells exhibit slow growth and increased UV sensitivity (59). The authors of this study concluded that DNA repair was more dependent on the poorly cooperative $(SSB)_{65}$ state whereas the highly cooperative $(SSB)_{35}$ state could be important for DNA replication (59). Although our present results pinpoint a role in DNA repair for the positive cooperative binding of plant Whirly proteins to ssDNA, both of these studies stress the biological importance of the cooperative binding of ssDNA-binding proteins to DNA. Another similarity between plant Whirlies and SSB proteins is the dependence of both higher order protein assembly and cooperative DNA binding on positively charged residues. Indeed, for the SSB proteins from *E. coli* and *T. aquaticus,* the mutation of an arginine residue to alanine results in a lower protein assembly and in a decrease in the cooperative binding to DNA (61). Overall, these studies outline additional similarities between plant Whirlies and SSB proteins although the mechanistic details differ between these two classes of proteins.

The 24-mer assembly of plant Whirly proteins closely resembles that of the *Homo sapiens*'s ferritin (62), that of the *Acidianus ambivalens*'s sulfur oxygenase reductase (63), that of the *Methanococcus jannaschii's* heat shock protein HSP16.5 (64) and that of the *E. coli*'s heat shock protein DegP (65). In each case, the assembly of 24 protomers generates a hollow protein shell with 432

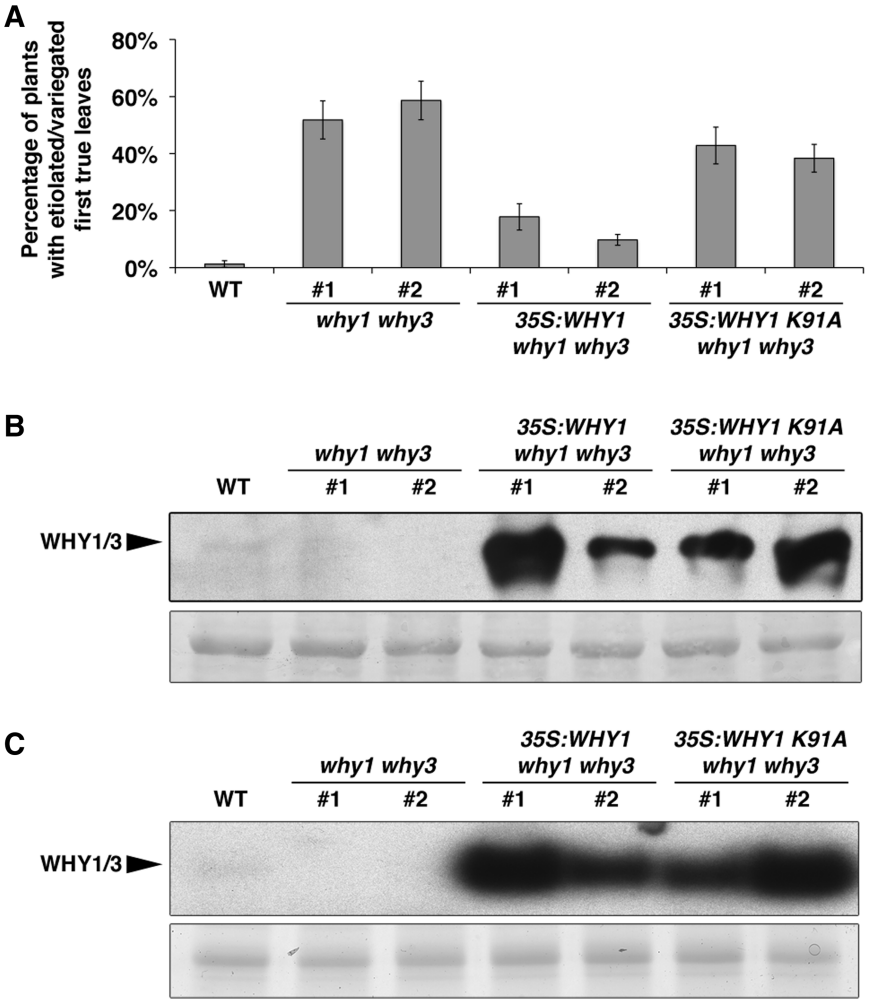

Figure 6. WHY1 but not WHY1 K91A partially complements ciprofloxacin sensitivity in *why1 why3* background. (**A**) Histogram displaying the percentage of plants with etiolated/variegated first true leaves after 10 days of growth on MS medium supplemented with 0.125 μM ciprofloxacin. For each genotype, 5 plates containing 50 plants on average were obtained. Error bars represent the standard deviation of the etiolation/variegation percentage measured on the different plates. For this experiment, the entire sequence of *WHY1* or *WHY1 K91A* under the control of the cauliflower mosaic virus 35S promoter was introduced into *why1 why3* plants. Except for the WT control, two independent homozygous lines were used for each genotype. A complete statistical treatment of these data can be found in Supplementary Table S1. (**B**) The protein level of WHY1/3 was assessed in each line by protein gel blot. Whirly proteins were detected by using an anti-WHY1/3 antibody. A section of the blot stained with Ponceau red and encompassing RbcL, the large subunit of Rubisco, is presented below as a loading control. (**C**) The ssDNA-binding activity of Whirly proteins was monitored by electrophoretic mobility shift assay using crude plastid protein extracts isolated from plants of the indicated genotypes and a radiolabeled dT_{32} oligonucleotide. A section of an SDS–PAGE stained with Coomassie blue is presented below as an equilibration control.

symmetry. The outer diameter of the shell varies from 12 nm for HSP16.5 to 19.5 nm for DegP. For all these nanostructures, pores or chimneys that lead to the interior of the sphere generally coincide with symmetry axes and Whirly proteins are no exception. A recent study on ferritin demonstrated that the mutation of single residues, termed oligomerization switch residues, is sufficient to abolish the formation of the spherical assembly (66). Our data indicates that K67 could fulfill a similar role for WHY2.

The peculiar assembly of plant Whirly proteins into 24-mers could promote binding to both DNA and metabolites. Indeed, most of the proteins that form nanostructures with 432 symmetry also store compounds. Ferritin is known to store iron atoms (62,67), the enzyme sulfur oxygenase reductase stores metabolites (63) whereas the heat shock proteins HSP16.5 and DegP both store proteins (68,69). One possibility would be that plant Whirly proteins bind and encapsulate harmful chemical compounds thereby protecting DNA against damage. In this regard, it is interesting that DPS, an homolog of ferritin that also forms a hollow protein shell, protects DNA from oxidative stress by binding DNA and by sequestering iron in its core (67). It is also worth mentioning that ferredoxin:sulfite reductase, an enzyme that localizes to chloroplast nucleoids, also binds concomitantly DNA and sulfites (70). It would therefore be possible that Whirly proteins protect DNA by sequestering harmful molecules.

In a recent report, Kwon *et al.* (71) confirmed that DNA DSBs are repaired in *Arabidopsis* chloroplasts through a microhomology-mediated DNA repair pathway

and brought mechanistic insights into the repair process. Of particular interest is the fact that a single DSB is repaired using microhomologies located more than 3 kb apart from each other. This suggests that during the repair process, very long stretches of ssDNA are unveiled before finding matching microhomologous DNA sequences. Through 24-mer assembly and positive cooperative binding, Whirly proteins could bind and protect these regions as well as the single-stranded regions which are abundantly present elsewhere in plant organelles genomes (72). This would enable faithful repair of DNA and prevent MMBIR product accumulation. Thus, our new findings complement and extend the model we previously proposed (15).

In conclusion, our data highlight a role of the second lysine of the KGKAAL motif of plant Whirly proteins in both the tolerance against DNA damages and higher order protein assembly. Further work is however required to determine precisely how the 24-mer assembly of plant Whirly proteins is involved in DNA repair. Also, it will be interesting to verify whether the 24-mer assembly of Whirly proteins is involved in metabolite binding and what is the functional consequence of this interaction.

ACCESSION NUMBERS

The atomic coordinates and structure factors for WHY2 K67A form I, WHY2 K67A form II and WHY2 K67A-dT_{32} (codes 3R9Y, 3R9Z and 3RA0, respectively) have been deposited in the Protein Data Bank, Research Collaboratory for Structural Bioinformatics, Rutgers University, New Brunswick, NJ (http://www.rcsb.org/).

SUPPLEMENTARY DATA

Supplementary Data are available at NAR Online.

ACKNOWLEDGEMENTS

The assistance of Jihyun Daniel Yi and Antonio Nanci with the AFM is gratefully acknowledged. We would also like to thank both beamlines personnel for their technical assistance.

FUNDING

Grants from the Natural Sciences and Engineering Research Council of Canada; Fonds québécois de la recherche sur la nature et les technologies (to both N.B. and J.S.); Canadian Institutes of Health Research (to J.S.); E.Z. and E.L. were funded by scholarships from the Natural Sciences and Engineering Research Council of Canada. Work was carried out in part at beamline X29 of the National Synchrotron Light Source. Financial support comes principally from the Offices of Biological and Environmental Research and of Basic Energy Sciences of the US Department of Energy, and from the National Center for Research Resources of the National Institutes of Health. Some of the research was performed at the 08ID-1 beamline at the Canadian Light Source. Funding for open access charge: Natural Sciences and Engineering Research Council of Canada.

Conflict of interest statement. None declared.

REFERENCES

1. Puchta,H. (2005) The repair of double-strand breaks in plants: mechanisms and consequences for genome evolution. *J. Exp. Bot.*, **56**, 1–14.
2. Kimura,S. and Sakaguchi,K. (2006) DNA repair in plants. *Chem. Rev.*, **106**, 753–766.
3. Nielsen,B.L., Cupp,J.D. and Brammer,J. (2010) Mechanisms for maintenance, replication, and repair of the chloroplast genome in plants. *J. Exp. Bot.*, **61**, 2535–2537.
4. Maréchal,A. and Brisson,N. (2010) Recombination and the maintenance of plant organelle genome stability. *New Phytol.*, **186**, 299–317.
5. Rhoads,D.M., Umbach,A.L., Subbaiah,C.C. and Siedow,J.N. (2006) Mitochondrial reactive oxygen species. Contribution to oxidative stress and interorganellar signaling. *Plant Physiol.*, **141**, 357–366.
6. Apel,K. and Hirt,H. (2004) Reactive oxygen species: metabolism, oxidative stress, and signal transduction. *Annu. Rev. Plant Biol.*, **55**, 373–399.
7. Palmer,J.D. (1983) Chloroplast DNA Exists in 2 Orientations. *Nature*, **301**, 92–93.
8. Lonsdale,D.M., Brears,T., Hodge,T.P., Melville,S.E. and Rottmann,W.H. (1988) The plant mitochondrial genome - homologous recombination as a mechanism for generating heterogeneity. *Philos. T. Roy. Soc. Lond. Ser. B Biol. Sci.*, **319**, 149–163.
9. Pang,Q., Hays,J.B. and Rajagopal,I. (1992) A plant cDNA that partially complements Escherichia coli recA mutations predicts a polypeptide not strongly homologous to RecA proteins. *Proc. Natl Acad. Sci. USA*, **89**, 8073–8077.
10. Cerutti,H., Osman,M., Grandoni,P. and Jagendorf,A.T. (1992) A homolog of Escherichia coli RecA protein in plastids of higher plants. *Proc. Natl Acad. Sci. USA*, **89**, 8068–8072.
11. Khazi,F.R., Edmondson,A.C. and Nielsen,B.L. (2003) An Arabidopsis homologue of bacterial RecA that complements an E. coli recA deletion is targeted to plant mitochondria. *Mol. Genet. Genomics.*, **269**, 454–463.
12. Cerutti,H. and Jagendorf,A.T. (1993) DNA strand-transfer activity in pea (Pisum sativum L.) chloroplasts. *Plant Physiol.*, **102**, 145–153.
13. Manchekar,M., Scissum-Gunn,K., Song,D., Khazi,F., McLean,S.L. and Nielsen,B.L. (2006) DNA recombination activity in soybean mitochondria. *J. Mol. Biol.*, **356**, 288–299.
14. Rowan,B.A., Oldenburg,D.J. and Bendich,A.J. (2010) RecA maintains the integrity of chloroplast DNA molecules in Arabidopsis. *J. Exp. Bot.*, **61**, 2575–2588.
15. Cappadocia,L., Maréchal,A., Parent,J.S., Lepage,E., Sygusch,J. and Brisson,N. (2010) Crystal structures of DNA-Whirly complexes and their role in Arabidopsis organelle genome repair. *Plant Cell*, **22**, 1849–1867.
16. Khakhlova,O. and Bock,R. (2006) Elimination of deleterious mutations in plastid genomes by gene conversion. *Plant J.*, **46**, 85–94.
17. Parent,J.S., Lepage,E. and Brisson,N. (2011) Divergent roles for the two PolI-like organelle DNA polymerases of Arabidopsis. *Plant Physiol.*, **156**, 254–262.
18. Hastings,P.J., Ira,G. and Lupski,J.R. (2009) A microhomology-mediated break-induced replication model for the origin of human copy number variation. *PLoS Genet.*, **5**, e1000327.
19. Desveaux,D., Maréchal,A. and Brisson,N. (2005) Whirly transcription factors: defense gene regulation and beyond. *Trends Plant Sci.*, **10**, 95–102.
20. Desveaux,D., Allard,J., Brisson,N. and Sygusch,J. (2002) A new family of plant transcription factors displays a novel ssDNA-binding surface. *Nat. Struct. Biol.*, **9**, 512–517.

21. Desveaux,D., Després,C., Joyeux,A., Subramaniam,R. and Brisson,N. (2000) PBF-2 is a novel single-stranded DNA binding factor implicated in PR-10a gene activation in potato. *Plant Cell*, **12**, 1477–1489.
22. Desveaux,D., Subramaniam,R., Després,C., Mess,J.N., Lévesque,C., Fobert,P.R., Dangl,J.L. and Brisson,N. (2004) A "Whirly" transcription factor is required for salicylic acid-dependent disease resistance in Arabidopsis. *Dev. Cell*, **6**, 229–240.
23. Yoo,H.H., Kwon,C., Lee,M.M. and Chung,I.K. (2007) Single-stranded DNA binding factor AtWHY1 modulates telomere length homeostasis in Arabidopsis. *Plant J.*, **49**, 442–451.
24. Xiong,J.Y., Lai,C.X., Qu,Z., Yang,X.Y., Qin,X.H. and Liu,G.Q. (2009) Recruitment of AtWHY1 and AtWHY3 by a distal element upstream of the kinesin gene AtKP1 to mediate transcriptional repression. *Plant Mol. Biol.*, **71**, 437–449.
25. Prikryl,J., Watkins,K.P., Friso,G., Wijk,K.J. and Barkan,A. (2008) A member of the Whirly family is a multifunctional RNA- and DNA-binding protein that is essential for chloroplast biogenesis. *Nucleic Acids Res.*, **36**, 5152–5165.
26. Melonek,J., Mulisch,M., Schmitz-Linneweber,C., Grabowski,E., Hensel,G. and Krupinska,K. (2010) Whirly1 in chloroplasts associates with intron containing RNAs and rarely co-localizes with nucleoids. *Planta.*, **232**, 471–481.
27. Krause,K., Kilbienski,I., Mulisch,M., Rödiger,A., Schäfer,A. and Krupinska,K. (2005) DNA-binding proteins of the Whirly family in Arabidopsis thaliana are targeted to the organelles. *FEBS Lett.*, **579**, 3707–3712.
28. Grabowski,E., Miao,Y., Mulisch,M. and Krupinska,K. (2008) Single-stranded DNA-binding protein Whirly1 in barley leaves is located in plastids and the nucleus of the same cell. *Plant Physiol.*, **147**, 1800–1804.
29. Maréchal,A., Parent,J.S., Véronneau-Lafortune,F., Joyeux,A., Lang,B.F. and Brisson,N. (2009) Whirly Proteins maintain plastid genome stability in Arabidopsis. *Proc. Natl Acad. Sci. USA*, **106**, 14693–14698.
30. Maréchal,A., Parent,J.S., Sabar,M., Véronneau-Lafortune,F., Abou-Rached,C. and Brisson,N. (2008) Overexpression of mtDNA-associated AtWhy2 compromises mitochondrial function. *BMC Plant Biol.*, **8**, 42.
31. Edmondson,A.C., Song,D., Alvarez,L.A., Wall,M.K., Almond,D., McClellan,D.A., Maxwell,A. and Nielsen,B.L. (2005) Characterization of a mitochondrially targeted single-stranded DNA-binding protein in Arabidopsis thaliana. *Mol. Genet. Genomics*, **273**, 115–122.
32. Horvath,M.P. (2008) In Rice,P.A. and Correll,C.C. (eds), *Protein-Nucleic Acid Interactions: Structural Biology*. Royal Society of Chemistry, London, pp. 91–128.
33. Kim,C. and Wold,M.S. (1995) Recombinant human replication protein A binds to polynucleotides with low cooperativity. *Biochemistry*, **34**, 2058–2064.
34. Lohman,T.M. and Ferrari,M.E. (1994) Escherichia coli single-stranded DNA-binding protein: multiple DNA-binding modes and cooperativities. *Annu. Rev. Biochem.*, **63**, 527–570.
35. Raghunathan,S., Kozlov,A.G., Lohman,T.M. and Waksman,G. (2000) Structure of the DNA binding domain of E. coli SSB bound to ssDNA. *Nat. Struct Biol.*, **7**, 648–652.
36. Cappadocia,L., Sygusch,J. and Brisson,N. (2008) Purification, crystallization and preliminary X-ray diffraction analysis of the Whirly domain of StWhy2 in complex with single-stranded DNA. *Acta. Crystallogr. Sect. F Struct. Biol. Cryst. Commun.*, **64**, 1056–1059.
37. Zheng,L., Baumann,U. and Reymond,J.L. (2004) An efficient one-step site-directed and site-saturation mutagenesis protocol. *Nucleic Acids Res.*, **32**, e115.
38. Emsley,P., Lohkamp,B., Scott,W.G. and Cowtan,K. Features and development of Coot. *Acta. Crystallogr. D Biol. Crystallogr.*, **66**, 486–501.
39. Brünger,A.T., Adams,P.D., Clore,G.M., DeLano,W.L., Gros,P., Grosse-Kunstleve,R.W., Jiang,J.S., Kuszewski,J., Nilges,M., Pannu,N.S. *et al.* (1998) Crystallography & NMR system: a new software suite for macromolecular structure determination. *Acta. Crystallogr. D Biol. Crystallogr.*, **54**, 905–921.
40. Brunger,A.T. (2007) Version 1.2 of the crystallography and NMR system. *Nat. Protoc.*, **2**, 2728–2733.
41. Adams,P.D., Afonine,P.V., Bunkoczi,G., Chen,V.B., Davis,I.W., Echols,N., Headd,J.J., Hung,L.W., Kapral,G.J., Grosse-Kunstleve,R.W. *et al.* PHENIX: a comprehensive Python-based system for macromolecular structure solution. *Acta. Crystallogr. D Biol. Crystallogr.*, **66**, 213–221.
42. DeLano,W.L. (2002) *The Pymol Molecular Graphics System*. DeLano Scientific, San Carlos, CA, USA.
43. Argaman,M., Golan,R., Thomson,N.H. and Hansma,H.G. (1997) Phase imaging of moving DNA molecules and DNA molecules replicated in the atomic force microscope. *Nucleic Acids Res.*, **25**, 4379–4384.
44. Hamon,L., Pastré,D., Dupaigne,P., Le Breton,C., Le Cam,E. and Piétrement,O. (2007) High-resolution AFM imaging of single-stranded DNA-binding (SSB) protein–DNA complexes. *Nucleic Acids Res.*, **35**, e58.
45. Hellens,R.P., Edwards,E.A., Leyland,N.R., Bean,S. and Mullineaux,P.M. (2000) pGreen: a versatile and flexible binary Ti vector for Agrobacterium-mediated plant transformation. *Plant Mol. Biol.*, **42**, 819–832.
46. Clough,S.J. and Bent,A.F. (1998) Floral dip: a simplified method for Agrobacterium-mediated transformation of Arabidopsis thaliana. *Plant J.*, **16**, 735–743.
47. Krissinel,E. and Henrick,K. (2007) Inference of macromolecular assemblies from crystalline state. *J. Mol. Biol.*, **372**, 774–797.
48. Lohman,T.M., Overman,L.B. and Datta,S. (1986) Salt-dependent changes in the DNA binding co-operativity of Escherichia coli single strand binding protein. *J. Mol. Biol.*, **187**, 603–615.
49. Reichel,C., Maas,C., Schulze,S., Schell,J. and Steinbiss,H.H. (1996) Cooperative binding to nucleic acids by barley yellow mosaic bymovirus coat protein and characterization of a nucleic acid-binding domain. *J. Gen. Virol.*, **77(Pt 4)**, 587–592.
50. Haseltine,C.A. and Kowalczykowski,S.C. (2002) A distinctive single-strand DNA-binding protein from the Archaeon Sulfolobus solfataricus. *Mol. Microbiol.*, **43**, 1505–1515.
51. Pfalz,J., Liere,K., Kandlbinder,A., Dietz,K.J. and Oelmüller,R. (2006) pTAC2, -6, and -12 are components of the transcriptionally active plastid chromosome that are required for plastid gene expression. *Plant Cell*, **18**, 176–197.
52. Shlyakhtenko,L.S., Lushnikov,A.Y., Li,M., Lackey,L., Harris,R.S. and Lyubchenko,Y.L. (2011) Atomic force microscopy studies provide direct evidence for dimerization of the HIV restriction factor APOBEC3G. *J. Biol. Chem.*, **286**, 3387–3395.
53. Luijsterburg,M.S., White,M.F., van Driel,R. and Dame,R.T. (2008) The major architects of chromatin: architectural proteins in bacteria, archaea and eukaryotes. *Crit. Rev. Biochem. Mol. Biol.*, **43**, 393–418.
54. Browning,D.F., Grainger,D.C. and Busby,S.J. (2010) Effects of nucleoid-associated proteins on bacterial chromosome structure and gene expression. *Curr. Opin. Microbiol.*, **13**, 773–780.
55. Cosgriff,S., Chintakayala,K., Chim,Y.T., Chen,X., Allen,S., Lovering,A.L. and Grainger,D.C. (2010) Dimerization and DNA-dependent aggregation of the Escherichia coli nucleoid protein and chaperone CbpA. *Mol. Microbiol.*, **77**, 1289–1300.
56. Ceci,P., Cellai,S., Falvo,E., Rivetti,C., Rossi,G.L. and Chiancone,E. (2004) DNA condensation and self-aggregation of Escherichia coli Dps are coupled phenomena related to the properties of the N-terminus. *Nucleic Acids Res.*, **32**, 5935–5944.
57. Mumtsidu,E., Makhov,A.M., Konarev,P.V., Svergun,D.I., Griffith,J.D. and Tucker,P.A. (2008) Structural features of the single-stranded DNA-binding protein of Epstein-Barr virus. *J. Struct. Biol.*, **161**, 172–187.
58. Meyer,R.R. and Laine,P.S. (1990) The single-stranded DNA-binding protein of Escherichia coli. *Microbiol. Rev.*, **54**, 342–380.
59. Carlini,L.E., Porter,R.D., Curth,U. and Urbanke,C. (1993) Viability and preliminary in vivo characterization of site-directed mutants of Escherichia coli single-stranded DNA-binding protein. *Mol. Microbiol.*, **10**, 1067–1075.
60. Ferrari,M.E., Fang,J. and Lohman,T.M. (1997) A mutation in E. coli SSB protein (W54S) alters intra-tetramer negative cooperativity and inter-tetramer positive cooperativity for single-stranded DNA binding. *Biophys. Chem.*, **64**, 235–251.

61. Witte,G., Fedorov,R. and Curth,U. (2008) Biophysical analysis of Thermus aquaticus single-stranded DNA binding protein. *Biophys. J.*, **94**, 2269–2279.
62. Theil,E.C. (1987) Ferritin: structure, gene regulation, and cellular function in animals, plants, and microorganisms. *Annu. Rev. Biochem.*, **56**, 289–315.
63. Urich,T., Gomes,C.M., Kletzin,A. and Frazão,C. (2006) X-ray structure of a self-compartmentalizing sulfur cycle metalloenzyme. *Science*, **311**, 996–1000.
64. Kim,K.K., Kim,R. and Kim,S.H. (1998) Crystal structure of a small heat-shock protein. *Nature*, **394**, 595–599.
65. Krojer,T., Sawa,J., Schafer,E., Saibil,H.R., Ehrmann,M. and Clausen,T. (2008) Structural basis for the regulated protease and chaperone function of DegP. *Nature*, **453**, 885–890.
66. Zhang,Y., Raudah,S., Teo,H., Teo,G.W., Fan,R., Sun,X. and Orner,B.P. (2010) Alanine-shaving mutagenesis to determine key interfacial residues governing the assembly of a nano-cage maxi-ferritin. *J. Biol. Chem.*, **285**, 12078–12086.
67. Grant,R.A., Filman,D.J., Finkel,S.E., Kolter,R. and Hogle,J.M. (1998) The crystal structure of Dps, a ferritin homolog that binds and protects DNA. *Nat. Struct. Biol.*, **5**, 294–303.
68. Krojer,T., Sawa,J., Schäfer,E., Saibil,H.R., Ehrmann,M. and Clausen,T. (2008) Structural basis for the regulated protease and chaperone function of DegP. *Nature*, **453**, 885–890.
69. Sutter,M., Boehringer,D., Gutmann,S., Günther,S., Prangishvili,D., Loessner,M.J., Stetter,K.O., Weber-Ban,E. and Ban,N. (2008) Structural basis of enzyme encapsulation into a bacterial nanocompartment. *Nat. Struct. Mol. Biol.*, **15**, 939–947.
70. Sekine,K., Fujiwara,M., Nakayama,M., Takao,T., Hase,T. and Sato,N. (2007) DNA binding and partial nucleoid localization of the chloroplast stromal enzyme ferredoxin:sulfite reductase. *FEBS J*, **274**, 2054–2069.
71. Kwon,T., Huq,E. and Herrin,D.L. (2010) Microhomology-mediated and nonhomologous repair of a double-strand break in the chloroplast genome of Arabidopsis. *Proc. Natl Acad. Sci. USA*, **107**, 13954–13959.
72. Backert,S., Lurz,R., Oyarzabal,O.A. and Börner,T. (1997) High content, size and distribution of single-stranded DNA in the mitochondria of Chenopodium album (L.). *Plant Mol. Biol.*, **33**, 1037–1050.
73. Davis,I.W., Leaver-Fay,A., Chen,V.B., Block,J.N., Kapral,G.J., Wang,X., Murray,L.W., Arendall,W.B. 3rd, Snoeyink,J., Richardson,J.S. *et al.* (2007) MolProbity: all-atom contacts and structure validation for proteins and nucleic acids. *Nucleic Acids Res.*, **35**, W375–W383.