Human Molecular Genetics, 2019, Vol. 28, No. 7 1090–1099

doi: 10.1093/hmg/ddy415
Advance Access Publication Date: 29 November 2018
General Article

GENERAL ARTICLE

Structural basis for adPEO-causing mutations in the mitochondrial TWINKLE helicase

Bradley Peter[1,*], Geraldine Farge[2], Carlos Pardo-Hernandez[1], Stefan Tångefjord[3] and Maria Falkenberg[1]

[1]Department of Medical Biochemistry and Cell Biology, University of Gothenburg, PO Box 440, Sweden, [2]Centre Nacionale de la Recherche Scientifique/Institut National de Physique Nucléaire et des Particules, Laboratoire de Physique de Clermont, Université Clermont Auvergne, BP 10448, F-63000 Clermont-Ferrand, France and [3]Discovery Sciences, IMED Biotech Unit, AstraZeneca, Gothenburg, Sweden

*To whom correspondence should be addressed at: Department of Medical Biochemistry and Cell Biology, University of Gothenburg, SE-405 30 Gothenburg, Sweden. Tel: +46317866702; Fax: +4631416108; Email: peter.bradley@gu.se

Abstract

TWINKLE is the helicase involved in replication and maintenance of mitochondrial DNA (mtDNA) in mammalian cells. Structurally, TWINKLE is closely related to the bacteriophage T7 gp4 protein and comprises a helicase and primase domain joined by a flexible linker region. Mutations in and around this linker region are responsible for autosomal dominant progressive external ophthalmoplegia (adPEO), a neuromuscular disorder associated with deletions in mtDNA. The underlying molecular basis of adPEO-causing mutations remains unclear, but defects in TWINKLE oligomerization are thought to play a major role. In this study, we have characterized these disease variants by single-particle electron microscopy and can link the diminished activities of the TWINKLE variants to altered oligomeric properties. Our results suggest that the mutations can be divided into those that (i) destroy the flexibility of the linker region, (ii) inhibit ring closure and (iii) change the number of subunits within a helicase ring. Furthermore, we demonstrate that wild-type TWINKLE undergoes large-scale conformational changes upon nucleoside triphosphate binding and that this ability is lost in the disease-causing variants. This represents a substantial advancement in the understanding of the molecular basis of adPEO and related pathologies and may aid in the development of future therapeutic strategies.

Introduction

Autosomal dominant progressive external ophthalmoplegia (adPEO) is a neurodegenerative and muscular disorder characterized by exercise intolerance, muscle weakness, peripheral neuropathy, deafness, ataxia and cataracts (1–3). Diagnosis is associated with mutations in genes required for mitochondrial DNA (mtDNA) maintenance and is linked to depletion or deletions of mtDNA in postmitotic tissues. Mutations in the TWINKLE helicase (encoded by the *TWNK* gene) as well as the two subunits of DNA polymerase gamma (encoded by the *POLG1* and *POLG2* genes) have been most closely linked with progression of the disease (4–6) (https://www.omim.org/entry/606075; https://www.omim.org/entry/157640).

TWINKLE is a ring-forming hexameric DNA helicase which localizes to mitochondrial nucleoids and is involved in adenosine triphosphate (ATP)-dependent unwinding of double-stranded DNA (dsDNA). TWINKLE activity is required for the

Received: October 16, 2018. **Revised:** November 27, 2018. **Accepted:** November 28, 2018

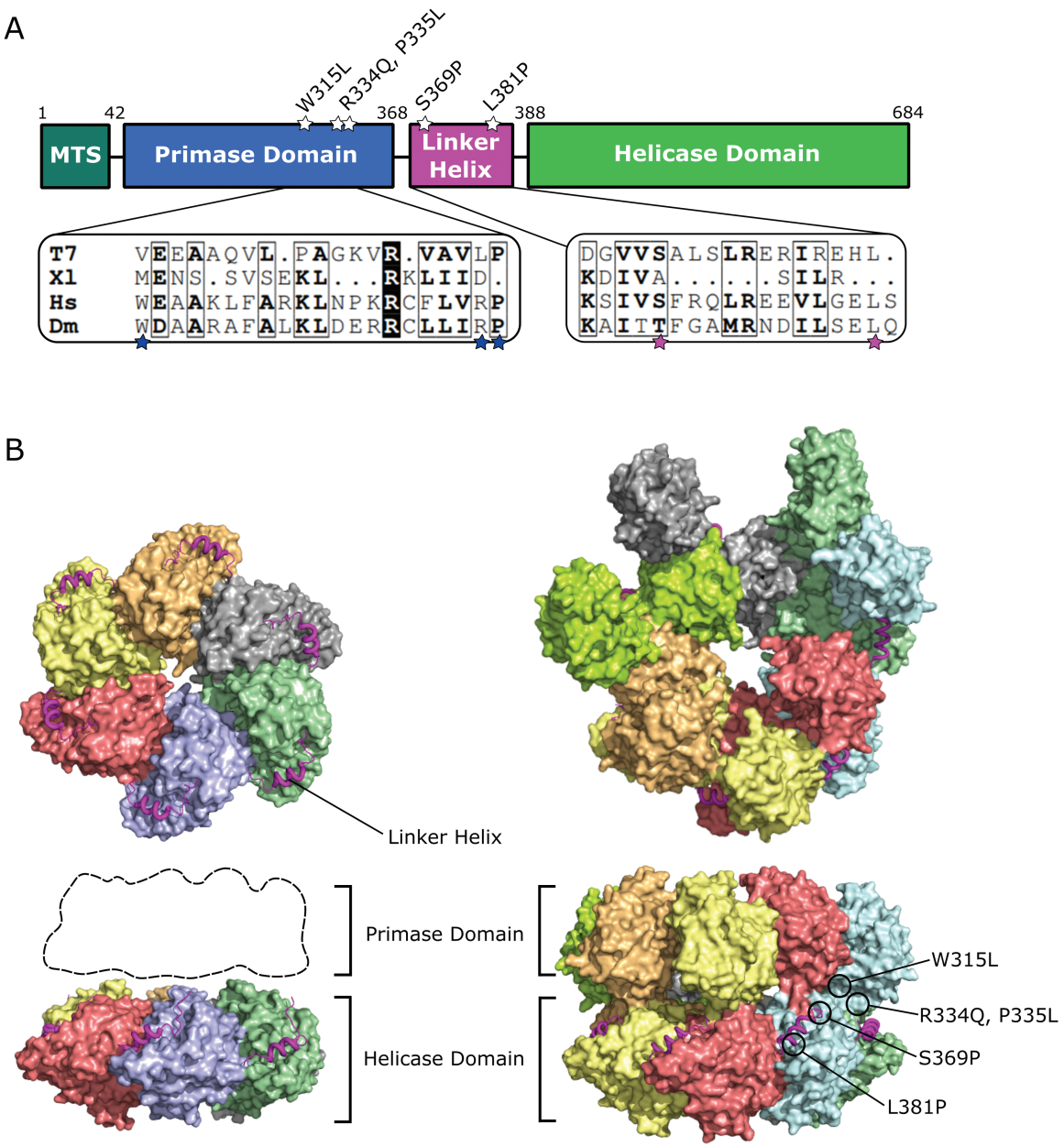

Figure 1. Domain organization of TWINKLE. (**A**) The domain architecture of human TWINKLE comprises a non-functional primase domain connected to a conserved helicase domain by a flexible linker helix. The positions of the disease-causing mutations analysed in this study are indicated. A sequence alignment of related proteins shows mild to strong conservation of residues in which point mutations result in disease (indicated by a star symbol); T7: bacteriophage T7 gp4; Xl: *Xenopus laevis* DNA2; Hs: *Homo sapiens* TWINKLE; Dm: *Drosophila melanogaster* mtDNA helicase. (**B**) Top and side views of the hexameric [left; protein data bank (PDB) ID: 1E0K] and heptameric (right; PDB ID: 5IKN) T7 gp4 protein. TWINKLE shares sequence and structure similarity to T7 gp4, making it an ideal model for comparison. Left: a fragment comprising the helicase domain and a part of the linker helix (residues 271–547) was crystallized as a symmetrical hexamer. Right: a double mutant lacking ATPase activity was crystallized as a heptamer (residues 64–549). The primase domain from one subunit is positioned above the helicase domain of a neighbouring subunit. In both structures, the linker helix (magenta) and domain organization are shown. The disease mutations studied here are all located either within the linker helix or in close proximity and are indicated in the heptameric structure.

replication and maintenance of human mtDNA together with the mtDNA polymerase (POLγ) and mitochondrial single-stranded DNA-binding protein (mtSSB) (4,7,8). TWINKLE belongs to the DnaB family and shares strong sequence and structure similarities with the bacteriophage T7 gp4 helicase (9,10). The domain layout of TWINKLE is similar to that of T7 gp4, comprising an N-terminal primase-like domain, a C-terminal helicase domain and a linker region connecting the two domains (Fig. 1A). There are no crystal structures available of TWINKLE but several structures of the homologous T7 gp4 as well as a low-resolution cryo-electron microscopy (EM) map of TWINKLE have been published (11–14). In these structures, both hexameric and heptameric rings are visible (see Fig. 1B for a schematic representation of the T7 gp4 structures). The linker helix forms a stable helix bundle at the surface of the helicase domain of the neighbouring subunit, causing the N-terminal primase domain to rest on top of the neighbouring helicase domain. The binding of nucleoside triphosphates (NTPs) occurs at the subunit

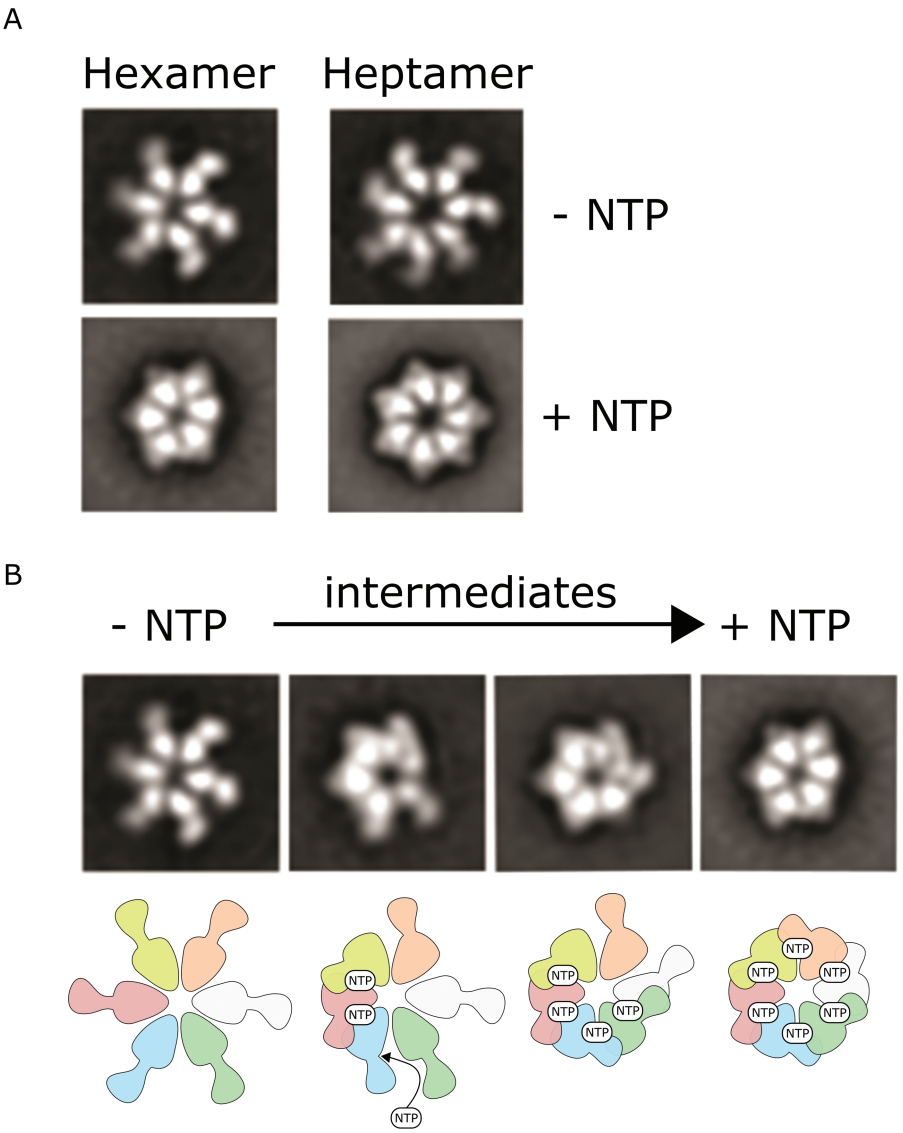

Figure 2. Wild-type TWINKLE undergoes nucleotide-dependent conformational changes. **(A)** TWINKLE forms hexameric and heptameric oligomers both in the absence and presence of nucleotides. In the absence of nucleotides, TWINKLE adopts a flexible star-like conformation with the primase domain appearing to extend out from the central pore via the linker domain. Upon nucleotide addition, the primase domain is repositioned above the neighbouring helicase domain resulting in a more compact structure. **(B)** A series of intermediate species were identified following nucleotide addition in which the primase domains were partially folded. This suggests that the folding of the primase domain may occur in a sequential manner. A schematic of this mechanism is presented, showing nucleotide binding at the subunit interface.

interface and, upon NTP hydrolysis, the helicase domains rotate and shift in relation to one another to provide the mechanical force required for DNA unwinding (12,15,16).

The linker region of T7 gp4 is crucial for both oligomerization and helicase activity (17,18) and the importance of this region is augmented by the fact that many of the TWINKLE mutations found in adPEO patients are positioned within the linker helix and its vicinity (19–22). Our group has previously taken advantage of the similarities in primary sequence that exist between TWINKLE and T7 gp4 to construct a 3D homology model of TWINKLE (21). In this model, the linker region interacts with both the primase domain and with the helicase domain of the neighbouring subunit. This suggests that mutations arising in the linker region may affect both intra-subunit folding as well as oligomerization (21). Previous studies by our group on the biochemical effects of these mutations have indeed shown that oligomer formation is affected and is coupled with severely reduced or non-existent catalytic activity (8,19,21). However, the exact nature of this oligomerization defect was not identified. The lack of structural data, coupled with the identification of further disease-causing mutations in TWINKLE necessitates a better understanding of the structural consequences of these mutations.

In this study, we performed a single-particle EM analysis of wild-type TWINKLE and the adPEO mutants W315L, R334Q, P335L, S369P and L381P in an effort to better understand the effect of these mutations on oligomerization and TWINKLE function. Our results strongly suggest that defects in oligomerization are responsible for the observed pathological changes in TWINKLE activity, with linker helix mutations disrupting subunit flexibility and primase domain mutations deregulating oligomer formation. We also observed nucleotide-dependent conformational changes in wild-type TWINKLE which were absent in disease-causing variants and may have pathological

Table 1. Distribution of oligomeric species of wild-type and mutant TWINKLE

	Oligomeric state (%)					
	Pentamer	Hexamer	Heptamer	Octamer	Nonamer	Broken ring
Wild-type	0	58.9	41.1	0	0	0
W315L	0	0	87.1	0	0	12.9
R334Q	0	36.6	55.1	0	0	8.3
P335L	0	4.7	72.3	14.8	1.8	6.4
S369P	0	13.2	41.6	0	0	45.2
L381P	11.5	33	28.6	0	0	26.9

The distribution of the various oligomeric states of wild-type and mutant TWINKLE are shown above. The values are given as a percentage of the number of each type of oligomer observed in all micrographs collected for each mutant and wild-type.

significance. This represents a substantial advancement in the understanding of the molecular basis of adPEO and related pathologies and may aid in the development of future therapeutic strategies.

Results

Wild-type TWINKLE undergoes nucleotide-dependent conformational changes

Our investigation of adPEO-causing TWINKLE mutants began with a negative stain analysis of the wild-type protein. A representative micrograph of negatively stained wild-type TWINKLE is shown in Supplementary Material, Figure S1. TWINKLE appeared almost exclusively in top-down views in our studies, thereby precluding 3D analysis of the helicase. Much like the homologous T7 gp4 helicase and previous TWINKLE studies (14,23), our reference-free 2D classifications showed a mixed population of both hexameric and heptameric species (Fig. 2A; Table 1). Interestingly, the domain organization of these oligomers differed depending on the presence or absence of NTPs. In the absence of NTPs, the oligomers formed star-like structures with the helicase domain forming an internal ring and the primase domain extending out into the solvent (Fig. 2A, upper panel). These domains are connected via the flexible linker helix. The diameter of the hexameric and heptameric species was ∼160 and 170 Å, respectively. The addition of NTPs induced both structural compaction and repositioning of the primase domain, forming hexamers and heptamers ∼120 and 130 Å in diameter, respectively (Fig. 2A, lower panel). By incubating wild-type TWINKLE with NTPs for 0–20 min, we were able to trap intermediates in the transition from the open NTP-free form to the compact NTP-bound form (Fig. 2B, upper panel). These 2D class averages show the primase domain rotated and repositioned on top of the neighbouring helicase domain. A simplified schematic of this process, in which the nucleotide-binding sites are located at the subunit interface, is shown in the lower panel of Figure 2B. This sequential folding mechanism suggests that the linker helix may interact with newly bound NTPs at the subunit interface to bring about repositioning of the primase domain.

The S369P and L381P linker helix mutations disrupt subunit flexibility

Two adPEO-causing mutations have been identified within the linker helix of TWINKLE: S369P and L381P. Ser369 is positioned at the beginning of the linker helix whereas Leu381 is located at the C-terminal end of the helix. In Figure 3A, the location of the equivalent residues in T7 gp4, Ser267 and Leu279 is shown. The linker helix rests on top of the neighbouring helicase domain, thus connecting one monomer to the next. Here, the linker helix forms extensive contacts with the neighbouring helicase domain, including numerous salt bridges and hydrogen bond interactions. It is likely that similar interactions persist in TWINKLE. In addition, both Ser267/Leu279 in T7 gp4 and Ser369/Leu381 in TWINKLE satisfy the requirements of helix-capping motifs (24–26). This suggests that these residues play important roles in stabilizing the linker helix.

Negative stain analysis of these mutants revealed that both S369P and L381P TWINKLE are capable of forming hexamers and heptamers, albeit at lower levels than wild-type (Fig. 3B and C; Table 1). In addition, a small population of L381P TWINKLE formed pentamers whereas changes in the number of monomers per oligomer were not observed for S369P TWINKLE. Interestingly, the large conformational changes of the primase domain observed upon NTP binding in the wild-type were not observed for either S369P or L381P TWINKLE (Supplementary Material, Fig. S2). In fact, no large difference between the class averages observed with or without NTPs was observed. This suggests a possible loss of flexibility due to the mutations. This significant reduction in subunit flexibility was perhaps the most notable effect of these mutations. Ring closure was severely undermined, with ∼45% of S369P and ∼27% of L381P TWINKLE existing as non-closed oligomers/broken rings (Fig. 3B and C; Table 1). This loss of flexibility was confirmed by differential scanning fluorimetry which revealed a thermal shift (ΔT_m) of + 4.2°C and + 2.5°C for S369P and L381P TWINKLE, respectively (Supplementary Material, Fig. S3). The introduction of a proline would likely restrain the possible conformations of the linker region, thus impairing normal ring closure as well as likely disrupting interactions between the linker helix and the neighbouring helicase domain. This is evident in biochemical analyses of the mutant proteins wherein S369P TWINKLE has severely reduced ATPase, helicase and single-stranded DNA (ssDNA)-binding activity and L381P TWINKLE is catalytically inactive (Supplementary Material, Table 1) (19,21).

Mutations in the TWINKLE primase domain target oligomerization

The role of the non-functional primase domain in TWINKLE is, as yet, unknown. However, three adPEO-causing mutations have been identified in this domain and in close proximity to the primase-linker boundary: W315L, R334Q and P335L. Because these residues are not strongly conserved in the T7 gp4 helicase (Fig. 1A), we have analysed the structure of a TWINKLE homology model previously published by our group (Fig. 4A) (21). In this

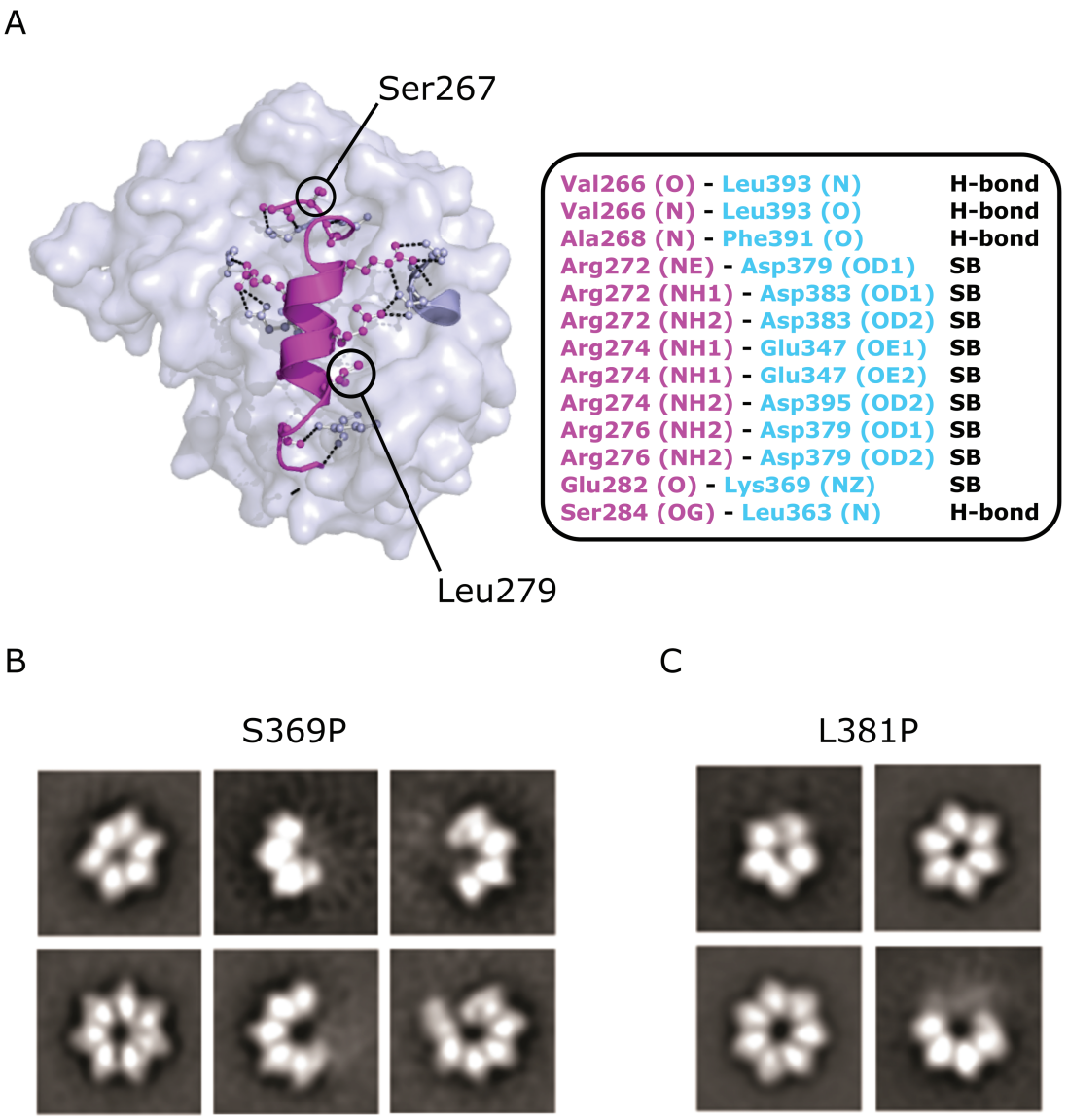

Figure 3. The S369P and L381P mutations in the linker helix of TWINKLE destabilize inter-subunit interactions. (**A**) The T7 gp4 linker helix (magenta) rests on top of the neighbouring helicase domain (cyan). The residues Ser267 and Leu279 in T7 gp4 correspond to Ser369 and Leu381in TWINKLE, respectively. The linker helix forms extensive contacts with the neighbouring helicase domain, including numerous salt bridges and hydrogen bond interactions. The Ser267 and Leu279 residues are located at/near the N- and C-termini of the helix and thus likely act to stabilize the helix. (**B**) S369P TWINKLE shows defects in ring closure. While some mature hexamers and heptamers are formed, a large fraction is incapable of forming closed (and therefore active) rings. (**C**) L381P TWINKLE is catalytically inactive and the 2D class averages in the presence and absence of nucleotides show that the hexamers and heptamers have lost their flexibility. In addition, the presence of pentamers and open oligomers suggests difficulties with oligomerization.

model, Arg334 and Pro335 are located on a loop adjacent to Trp315, with Arg334 forming a putative salt bridge with Asp310. All three residues are located within a conserved electropositive cleft thought to be crucial for DNA binding (21). Electrophoretic mobility shift assay (EMSA) analysis showed that binding of TWINKLE to both ssDNA and dsDNA was not affected by the R334Q and P335L mutations (Fig. 4E; Supplementary Material, Fig. S4). In contrast, the W315L mutant showed severe difficulty in binding to ssDNA while having wild-type like dsDNA-binding capacity. Negative stain studies of the W315L mutant reveal an overwhelming preference for heptamers (~87%), with the remainder having difficulties forming closed rings (Fig. 4B; Table 1). In previous enzymatic assays, this mutant showed dramatically reduced ssDNA binding, ATPase and helicase activity, with little observable activity in DNA replication assays (Supplementary Material, Table 1) (19,21). The R334Q mutant forms both hexamers and heptamers in ratios similar to wild-type (Fig. 4C; Table 1). However, the primase domain appears diffuse and disorganized, yielding 2D class averages with significant heterogeneity. Our previous biochemical analyses suggest that this mutant is only mildly disruptive, with ATPase and replication activity still possible albeit at lower than wild-type levels (Supplementary Material, Table 1) (19,21). The third mutant, P335L, displayed clear structural defects, existing in a variety of oligomeric states (Fig. 4D; Table 1). This deregulation of oligomerization led to the formation of higher-order octomers and nonamers in addition to the predominant heptamers (72%) and hexamers (5%). Biochemically, this mutant

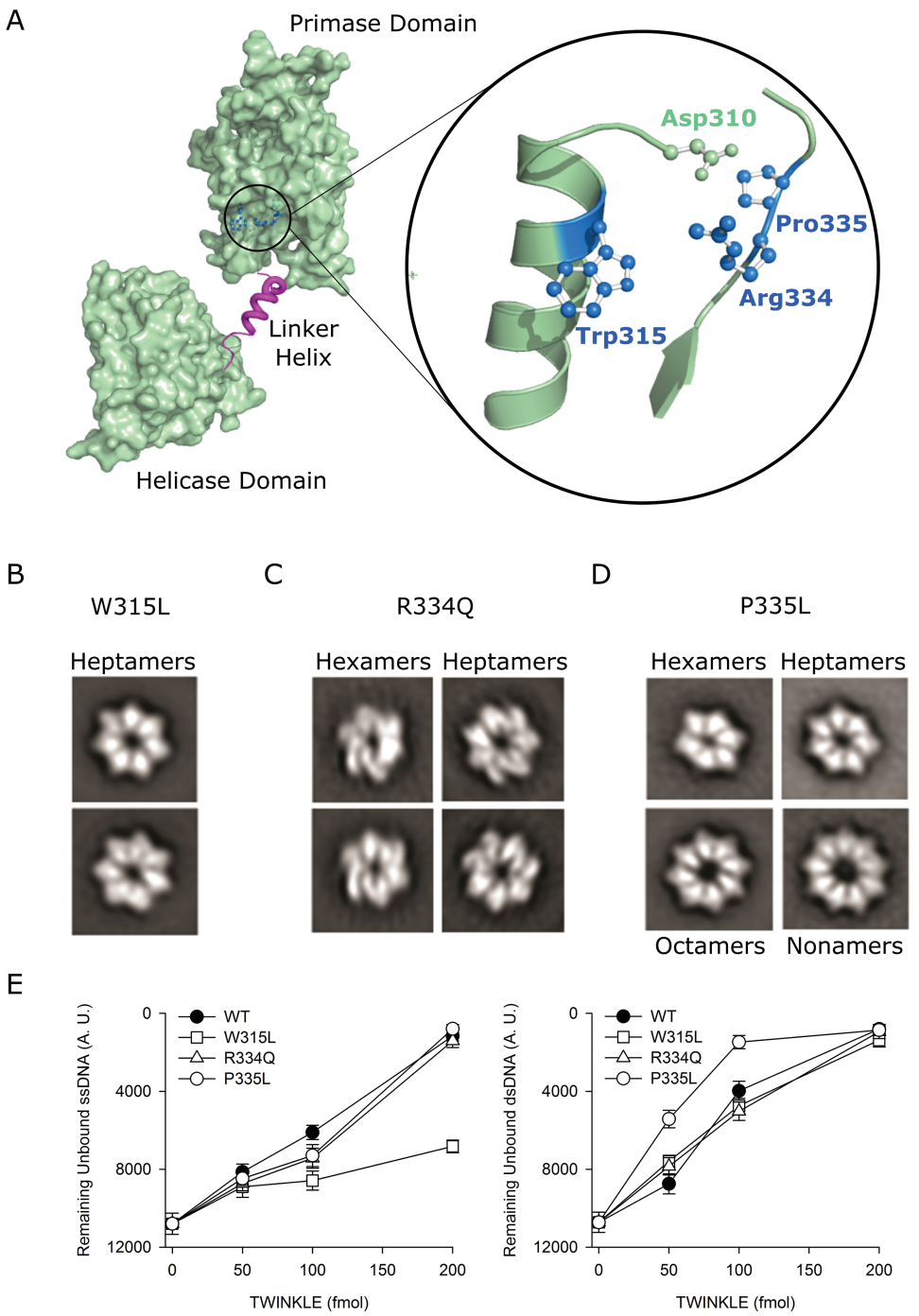

Figure 4. Mutations in the primase domain of TWINKLE affect both oligomer flexibility and degree of oligomerization. (**A**) A homology model of a single TWINKLE monomer has previously been constructed based on the T7 gp4 protein (21). Arg334 and Pro335 are located on a loop adjacent to Trp315, with Arg334 forming a putative salt bridge with Asp310. All three residues are located within a conserved electropositive cleft thought to be crucial for DNA binding. These residues are also likely in close proximity to the neighbouring primase domain in the hexameric/heptameric conformation. (**B**) The W315L mutant is capable only of forming heptamers, with ~12% not being able to form closed rings. (**C**) The R334Q mutant forms both hexamers and heptamers in ratios similar to wild-type. However, the structures are heterogeneous due in part to an apparent increase in primase domain flexibility. (**D**) The P335L mutant adopts a variety of oligomeric states, ranging from hexamers and heptamers to octomers and nonamers. (**E**) EMSA analysis shows that binding of TWINKLE to both ssDNA and dsDNA is not affected by the R334Q and P335L mutations, whereas W315L TWINKLE has difficulty binding to ssDNA while dsDNA-binding capacity is preserved.

showed poor ATPase and helicase activity and was poorly active in replication assays (Supplementary Material, Table 1) (19,21).

Much like the linker helix mutants, all of the primase domain mutants lacked the characteristic NTP-induced conformational change seen in the wild-type (Supplementary Material, Fig. S2). A similar trend was also observed in thermal stability studies, with all three mutants displaying increased stability relative to the wild-type (ΔT_m of +1.5°C, +3.6°C and +3.1°C for W315L, R334Q and P335L, respectively) (Supplementary Material, Fig. S3).

Discussion

Mutations in the TWINKLE helicase have previously been linked to severe mitochondrial pathologies, most notably adPEO (1,2). Subsequent studies by our group and others have attempted to characterize the biochemical defects of these variants in an effort to understand the underlying molecular basis of adPEO (19,21). Here we present evidence that strongly suggests defects in oligomerization are responsible for the observed pathological changes in TWINKLE activity. Furthermore, we show that wild-type TWINKLE undergoes nucleotide-dependent conformational changes which are absent in disease-causing variants.

Unlike the T7 gp4 helicase, TWINKLE can form oligomers in the absence of nucleotides or DNA (23,27). Interestingly, we observed that nucleotide binding induces a sequential conformational change wherein the primase domain folds over the helicase domain of the neighbouring subunit (Fig. 2). The linker region of one monomer forms extensive contacts with the adjacent helicase domain, connecting the subunits in a tandem fashion. Both crystal structures of the T7 gp4 helicase and EM reconstructions of TWINKLE support our observation (12,14). In the T7 gp4 helicase, deoxythymidine triphosphate binds at the subunit interface and in close proximity to the linker helix (12) and is thought to bind to a similar location in TWINKLE (21). *In vivo*, the NTP concentration is ∼3 mM and the open, star-like form of TWINKLE that we observed probably never exists. It does, however, illustrate the large conformational changes that occur upon nucleotide binding and release within one round of hydrolysis through the helicase ring. Ring flexibility is important not only for loading onto DNA but also for executing successive rounds of DNA unwinding. Both T7 gp4 and the RecA family of helicases require correct oligomerization and flexibility within the linker region for NTP hydrolysis and function (18,28–30). The substantial intra-subunit movements that take place during one round of DNA unwinding put high demands on this region linking one monomer to the next. It has to be very flexible, while still forming a stable interaction with the neighbouring subunit. This conformational flexibility was noticeably absent in the TWINKLE disease variants studied (Supplementary Material, Fig. S2).

Two of the mutations studied here, L381P and S369P, are located within the linker region and both have dramatic effects on the oligomeric properties of TWINKLE. The L381P mutation is located near a flexible loop close to the body of the helicase domain, a region that shows the largest conformational variation within the various T7 crystal structures (11–13). This variant is catalytically inactive and the 2D class averages in the presence and absence of nucleotides show that the hexamers and heptamers have lost their flexibility. In addition, the presence of pentamers (11.7%) and open oligomers (26.9%) suggests difficulties with oligomerization. The S369P mutation is positioned at the opposite end of the linker helix, close to the primase domain, and in this case ring closure appears to be affected. While some mature hexamers and heptamers are formed (13.2% and 41.6%, respectively), a large fraction (45.2%) of S369P TWINKLE is incapable of forming closed (and therefore active) rings. The introduction of a proline, which restricts flexibility by only adopting one of two conformations, seems the most likely reason for ring breakage. This loss of rotational freedom of the individual helicase subunits in relation to each other would likely prevent both NTP hydrolysis and the subsequent propelling of DNA by the DNA-binding loops within the central channel. Indeed, we observed that mutant TWINKLE variants had severe ATPase defects (Supplementary Material, Table 1). This may help to explain why many of these mutations exhibit a dominant effect in adPEO patients. Since TWINKLE activity is dependent on the continuous cycling through individual subunits within the ring, the presence of a single inactive mutant subunit would be sufficient to stall the catalytic cycle. We have attempted to recreate this situation *in vitro* by mixing the wild-type and mutant proteins; however, we could not consistently control the dissociation and reformation of heterogeneous subunits, and this presents an opportunity for future work. Analysis of the biochemistry of these mixtures is, however, possible and this has been performed previously for a number of adPEO TWINKLE variants both *in vitro* (19) and *in vivo* (31). These studies showed that the addition of a wild-type copy of TWINKLE can partially stabilize hexamerization but that this does not rescue helicase activity. This is likely a result of improper alignment of wild-type and mutant subunits which is required to transmit the energy from ATP-dependent conformational rearrangements to DNA unwinding. As we have shown, the subunit interface comprises a large network of interactions which, when disturbed, can have dramatic structural and functional effects. Thus, the presence of even a single mutant subunit will likely disturb helicase activity, even if one or more wild-type subunits are present. This is especially valid for the severe linker mutations S369P and L381P. Other milder mutations, such as W315L and R334Q, may be more difficult to analyse *in vivo*. In support of this notion, an A360T mutant which displayed only a mild reduction in mtDNA synthesis *in vitro* did not cause any noticeable phenotype upon overexpression in mice (32).

These variants associated with milder pathophysiology (W315L, R334Q and P335L) displayed less dramatic effects on ring flexibility (<13% broken rings) and are located within the primase domain. Studies of prokaryotic primases have identified a conserved region of electropositive potential within the catalytic core that is required for ssDNA binding (33). The W315L, R334Q and P335L mutations are all located within this region. Although TWINKLE has lost its primase function, this region is still thought to be involved in ssDNA binding (34). In agreement with this observation, W315L TWINKLE displayed reduced ssDNA-binding capacity (Fig. 4E; Supplementary Material, Fig. S4).

Although these mutations did not appear to cause significant ring breakage, an effect on oligomerization was still seen. More specifically, the W315L variant existed almost exclusively as a heptamer (87.1%) whereas the P335L variant formed higher-order oligomers (likely a result of reduced conformational strain imposed by Pro335). The presence of both hexameric and heptameric species of wild-type TWINKLE raises the question of which oligomer represents active TWINKLE and whether shifts in these populations can have a negative effect. Previous studies have speculated that ejection of one subunit from a heptameric helicase may provide a mechanism by which to open the helicase ring and load a hexamer onto ssDNA (35) while others suggest that T7 gp4 heptamers are processive on dsDNA (13). In contrast to T7 gp4, TWINKLE binds dsDNA with a much higher affinity than ssDNA (27,34). Crystal structures of T7 gp4 as well as EM models of TWINKLE show that the central channel of the hexamer is wide enough to accommodate ssDNA but not dsDNA (12,14). This suggests that both hexameric and heptameric TWINKLE are required in order to achieve efficient DNA loading and unwinding. Disruption of the hexamer-to-heptamer ratio may therefore have pathological consequences, such was observed for the W315L TWINKLE variant (Table 1).

It is also important to consider the spatial arrangement and interactions of the TWINKLE helicase with the other

components of the mtDNA replisome, namely mtSSB and POLγ. Rolling circle DNA replication assays combining wild-type or mutant TWINKLE with mtSSB and POLγ suggest that adPEO mutations have a strong negative effect on these protein–protein interactions (Supplementary Material, Table 1). Both crystal structures and cryo-EM maps of the T7 replisome have recently been solved (36,37). In these structures, the helicase domain appears to form the majority of contacts with the DNA polymerase (Supplementary Material, Fig. S5). With the exception of L381P, the majority of adPEO mutations lie outside of the helicase–polymerase interface (Supplementary Material, Fig. S5B). Of course, given the effect of these mutations on ring flexibility and stability, it is likely that the helicase–polymerase interaction will be disturbed. It should also be noted that although TWINKLE and POLγ share a high sequence and structure similarity with their T7 counterparts, differences in the way in which DNA is replicated in mitochondria and T7 bacteriophages (e.g. coupling of leading and lagging strand replication) may mean that the effect of these mutations cannot be accurately modelled on the T7 replisome. It has also been speculated that TWINKLE directly interacts with mtSSB (7). Physical interactions between replicative helicases and ssDNA-binding proteins have been observed in other systems (38–40), providing yet another possible mechanism by which adPEO mutations in TWINKLE may cause malfunction and disease.

Mutations in the TWINKLE helicase are also responsible for a wide range of disorders in addition to adPEO. Perrault disease (R391H, W441G, V507I and N585S), hepatocerebral mtDNA depletion syndrome/MDS (A318T, T457I and Y508C) and infantile-onset spinocerebellar ataxia/IOSCA (Y508C) are all caused by recessively inherited mutations in TWINKLE (41–44). It is possible that defects in TWINKLE oligomerization may also be responsible for these pathologies. As an example, Y508C, implicated in both MDS and IOSCA, forms a pocket for Ile367 in the linker region on the neighbouring subunit and may therefore hinder oligomerization. As more disease-causing TWINKLE variants are discovered, the need for high resolution structural information becomes greater. Future projects will attempt to solve the structure of TWINKLE and the mitochondrial replisome at high resolution by cryo-EM in an effort to more fully understand TWINKLE pathologies as well as attempt to identify therapeutic targets.

In summary, our results provide strong evidence that defects in oligomerization resulting from mutations in the linker helix and primase domain of TWINKLE are a primary cause of the development of adPEO and potentially other life-threatening mitochondrial pathologies. Furthermore, we highlight the dynamic nature of the structure of TWINKLE and the coupling of nucleotide binding to conformational change and ATP hydrolysis.

Materials and Methods

Recombinant protein expression and purification

Wild-type TWINKLE lacking the mitochondrial targeting sequence (residues 1–42) was expressed in *Sf*9 insect cells and purified as previously described (7). Single amino acid substitutions were introduced using a QuickChange Lightning (Stratagene, San Diego, CA) site-directed mutagenesis kit according to the manufacturer's instructions. Mutant TWINKLE variants were expressed and purified as for wild-type. Purified proteins were dialysed and stored in an analysis buffer [25 mM Tris–HCl pH 7.5, 10 mM $MgCl_2$, 1 mM DTT and 10% (v/v) glycerol].

EMSA

Binding of TWINKLE to both ssDNA and dsDNA was assayed essentially as previously described (34). Briefly, a 30-mer poly-dT oligonucleotide was 5′ end-labelled with [γ-^{32}P] ATP using T4 polynucleotide kinase (New England Biolabs, Ipswich, MA) to form the ssDNA substrate which was subsequently annealed to a 30-mer poly-dA oligonucleotide to form the dsDNA substrate. The reaction mixture (15 µl) contained 20 fmol DNA template, 20 mM Tris–HCl (pH 8.0), 1 mM DTT, 0.1 mg/ml bovine serum albumin, 10 mM $MgCl_2$, 10% glycerol and 2 mM ATP. Increasing amounts of TWINKLE (0, 50, 100 and 200 fmol, calculated as hexamer) were added and the samples were incubated at room temperature for 10 min. DNA binding was analysed by electrophoresis using a 6% polyacrylamide gel run at 100 V for 15 min. The band corresponding to free DNA template was quantified using phosphorimager and MultiGauge software (Fujifilm, Tokyo, Japan). All experiments were performed in triplicate and representative results shown.

Grid preparation

Wild-type and mutant TWINKLE were diluted to 0.05 mg/ml in analysis buffer in the absence or presence of 4 mM uridine triphosphate and incubated at 20°C for 20 min. Samples (5 µl) were incubated on freshly glow-discharged 400 mesh carbon-coated copper grids (Axlab, Vedbæk, Denmark) for 1 min, after which excess liquid was removed using filter paper. The grids were washed once each in ddH_2O and 0.75% (w/v) uranyl formate followed by 30 s incubation on a drop of stain. Excess stain was removed and the grids were allowed to air-dry overnight.

Data collection and analysis

Negative stain images were collected at a nominal magnification of 100 000× on a 120 kV LEO 912 microscope equipped with a 2k × 2k VELETA Olympus CCD camera. An omega filter was used to remove inelastically scattered electrons. Images were recorded at a defocus of 0.6 µm. At this defocus, the first zero of the CTF was positioned at ∼20 Å. Particles were boxed using the e2boxer2.py script from the EMAN2 suite (45) in swarm mode using a box size of 50 × 50 pixels. The particles were edge-normalized and classified using the reference-free maximum likelihood program mlf_align2d in XMIPP (46–49). The first round of classification was performed using a high-resolution limit of 30 Å and a coarse psi angle interval. The second round of classification was done using a high-resolution limit set to 20 Å and a finer psi angle search interval. The use of an additional norm parameter, which refines the normalization of the images internally, significantly improved the final class averages and reduced the overfitting of noise in the images (50).

Supplementary Material

Supplementary Material is available at *HMG* online.

Conflict of Interest statement. None declared.

Funding

Swedish Research Council (VR521-2013-3621 to M.F.); Swedish Cancer Foundation (CAN 2016/816 to M.F.); European Research Council (683191 to M.F.); Inga Britt and Arne Lundberg Foundation; Knut and Alice Wallenbergs Foundation (KAW 2011

to M.F., KAW 2014 to M.F.); Association Française Contre les Myopathies Téléthon (21411 to G.F.). We also acknowledge the Centre for Cellular Imaging at the University of Gothenburg and the National Microscopy Infrastructure (VR-RFI 2016-00968) for providing assistance in microscopy.

References

1. Zeviani, M., Servidei, S., Gellera, C., Bertini, E., DiMauro, S. and DiDonato, S. (1989) An autosomal dominant disorder with multiple deletions of mitochondrial DNA starting at the D-loop region. *Nature*, **339**, 309–311.
2. Suomalainen, A., Majander, A., Wallin, M., Setala, K., Kontula, K., Leinonen, H., Salmi, T., Paetau, A., Haltia, M., Valanne, L. *et al.* (1997) Autosomal dominant progressive external ophthalmoplegia with multiple deletions of mtDNA: clinical, biochemical, and molecular genetic features of the 10q-linked disease. *Neurology*, **48**, 1244–1253.
3. Hirano, M., Davidson, M. and DiMauro, S. (2001) Mitochondria and the heart. *Curr. Opin. Cardiol.*, **16**, 201–210.
4. Spelbrink, J.N., Li, F.Y., Tiranti, V., Nikali, K., Yuan, Q.P., Tariq, M., Wanrooij, S., Garrido, N., Comi, G., Morandi, L. *et al.* (2001) Human mitochondrial DNA deletions associated with mutations in the gene encoding Twinkle, a phage T7 gene 4-like protein localized in mitochondria. *Nat. Genet.*, **28**, 223–231.
5. Van Goethem, G., Dermaut, B., Lofgren, A., Martin, J.J. and Van Broeckhoven, C. (2001) Mutation of POLG is associated with progressive external ophthalmoplegia characterized by mtDNA deletions. *Nat. Genet.*, **28**, 211–212.
6. Longley, M.J., Clark, S., Yu Wai Man, C., Hudson, G., Durham, S.E., Taylor, R.W., Nightingale, S., Turnbull, D.M., Copeland, W.C. and Chinnery, P.F. (2006) Mutant POLG2 disrupts DNA polymerase gamma subunits and causes progressive external ophthalmoplegia. *Am. J. Hum. Genet.*, **78**, 1026–1034.
7. Korhonen, J.A., Gaspari, M. and Falkenberg, M. (2003) TWINKLE has 5' → 3' DNA helicase activity and is specifically stimulated by mitochondrial single-stranded DNA-binding protein. *J. Biol. Chem.*, **278**, 48627–48632.
8. Korhonen, J.A., Pham, X.H., Pellegrini, M. and Falkenberg, M. (2004) Reconstitution of a minimal mtDNA replisome in vitro. *EMBO J.*, **23**, 2423–2429.
9. Leipe, D.D., Aravind, L., Grishin, N.V. and Koonin, E.V. (2000) The bacterial replicative helicase DnaB evolved from a RecA duplication. *Genome Res.*, **10**, 5–16.
10. Shutt, T.E. and Gray, M.W. (2006) Twinkle, the mitochondrial replicative DNA helicase, is widespread in the eukaryotic radiation and may also be the mitochondrial DNA primase in most eukaryotes. *J. Mol. Evol.*, **62**, 588–599.
11. Sawaya, M.R., Guo, S., Tabor, S., Richardson, C.C. and Ellenberger, T. (1999) Crystal structure of the helicase domain from the replicative helicase–primase of bacteriophage T7. *Cell*, **99**, 167–177.
12. Singleton, M.R., Sawaya, M.R., Ellenberger, T. and Wigley, D.B. (2000) Crystal structure of T7 gene 4 ring helicase indicates a mechanism for sequential hydrolysis of nucleotides. *Cell*, **101**, 589–600.
13. Toth, E.A., Li, Y., Sawaya, M.R., Cheng, Y. and Ellenberger, T. (2003) The crystal structure of the bifunctional primase–helicase of bacteriophage T7. *Mol. Cell*, **12**, 1113–1123.
14. Fernandez-Millan, P., Lazaro, M., Cansiz-Arda, S., Gerhold, J.M., Rajala, N., Schmitz, C.A., Silva-Espina, C., Gil, D., Bernado, P., Valle, M. *et al.* (2015) The hexameric structure of the human mitochondrial replicative helicase Twinkle. *Nucleic Acids Res.*, **43**, 4284–4295.
15. Crampton, D.J., Guo, S., Johnson, D.E. and Richardson, C.C. (2004) The arginine finger of bacteriophage T7 gene 4 helicase: role in energy coupling. *Proc. Natl. Acad. Sci. U. S. A.*, **101**, 4373–4378.
16. Donmez, I. and Patel, S.S. (2006) Mechanisms of a ring shaped helicase. *Nucleic Acids Res.*, **34**, 4216–4224.
17. Guo, S., Tabor, S. and Richardson, C.C. (1999) The linker region between the helicase and primase domains of the bacteriophage T7 gene 4 protein is critical for hexamer formation. *J. Biol. Chem.*, **274**, 30303–30309.
18. Lee, S.J. and Richardson, C.C. (2004) The linker region between the helicase and primase domains of the gene 4 protein of bacteriophage T7. Role in helicase conformation and activity. *J. Biol. Chem.*, **279**, 23384–23393.
19. Korhonen, J.A., Pande, V., Holmlund, T., Farge, G., Pham, X.H., Nilsson, L. and Falkenberg, M. (2008) Structure–function defects of the TWINKLE linker region in progressive external ophthalmoplegia. *J. Mol. Biol.*, **377**, 691–705.
20. Goffart, S., Cooper, H.M., Tyynismaa, H., Wanrooij, S., Suomalainen, A. and Spelbrink, J.N. (2009) Twinkle mutations associated with autosomal dominant progressive external ophthalmoplegia lead to impaired helicase function and in vivo mtDNA replication stalling. *Hum. Mol. Genet.*, **18**, 328–340.
21. Holmlund, T., Farge, G., Pande, V., Korhonen, J., Nilsson, L. and Falkenberg, M. (2009) Structure–function defects of the twinkle amino-terminal region in progressive external ophthalmoplegia. *Biochim. Biophys. Acta*, **1792**, 132–139.
22. Echaniz-Laguna, A., Chanson, J.B., Wilhelm, J.M., Sellal, F., Mayencon, M., Mohr, M., Tranchant, C. and Mousson de Camaret, B. (2010) A novel variation in the Twinkle linker region causing late-onset dementia. *Neurogenetics*, **11**, 21–25.
23. Crampton, D.J., Ohi, M., Qimron, U., Walz, T. and Richardson, C.C. (2006) Oligomeric states of bacteriophage T7 gene 4 primase/helicase. *J. Mol. Biol.*, **360**, 667–677.
24. Presta, L.G. and Rose, G.D. (1988) Helix signals in proteins. *Science*, **240**, 1632–1641.
25. Richardson, J.S. and Richardson, D.C. (1988) Helix lap-joints as ion-binding sites: DNA-binding motifs and Ca-binding "EF hands" are related by charge and sequence reversal. *Proteins*, **4**, 229–239.
26. Sagermann, M., Martensson, L.G., Baase, W.A. and Matthews, B.W. (2002) A test of proposed rules for helix capping: implications for protein design. *Protein Sci.*, **11**, 516–521.
27. Jemt, E., Farge, G., Backstrom, S., Holmlund, T., Gustafsson, C.M. and Falkenberg, M. (2011) The mitochondrial DNA helicase TWINKLE can assemble on a closed circular template and support initiation of DNA synthesis. *Nucleic Acids Res.*, **39**, 9238–9249.
28. Ariza, A., Richard, D.J., White, M.F. and Bond, C.S. (2005) Conformational flexibility revealed by the crystal structure of a crenarchaeal RadA. *Nucleic Acids Res.*, **33**, 1465–1473.
29. Bailey, S., Eliason, W.K. and Steitz, T.A. (2007) Structure of hexameric DnaB helicase and its complex with a domain of DnaG primase. *Science*, **318**, 459–463.
30. Zhang, Y., Palla, M., Sun, A. and Liao, J.C. (2013) Identification of unique interactions between the flexible linker and the RecA-like domains of DEAD-box helicase Mss116. *J. Phys. Condens. Matter*, **25**, 374101.

31. Matsushima, Y. and Kaguni, L.S. (2007) Differential phenotypes of active site and human autosomal dominant progressive external opthalmoplegia mutations in *Drosophila* mitochondrial DNA helicase expressed in Schneider cells. *J. Biol. Chem.*, **282**, 9436–9444.
32. Tyynismaa, H., Mjosund, K.P., Wanrooij, S., Lappalainen, I., Ylikallio, E., Jalanko, A., Spelbrink, J.N., Paetau, A. and Suomalainen, A. (2005) Mutant mitochondrial helicase Twinkle causes multiple mtDNA deletions and a late-onset mitochondrial disease in mice. *Proc. Natl. Acad. Sci. U. S. A.*, **102**, 17687–17692.
33. Lee, S.J. and Richardson, C.C. (2005) Acidic residues in the nucleotide-binding site of the bacteriophage T7 DNA primase. *J. Biol. Chem.*, **280**, 26984–26991.
34. Farge, G., Holmlund, T., Khvorostova, J., Rofougaran, R., Hofer, A. and Falkenberg, M. (2008) The N-terminal domain of TWINKLE contributes to single-stranded DNA binding and DNA helicase activities. *Nucleic Acids Res.*, **36**, 393–403.
35. O'Shea, V.L. and Berger, J.M. (2014) Loading strategies of ring-shaped nucleic acid translocases and helicases. *Curr. Opin. Struct. Biol.*, **25**, 16–24.
36. Kulczyk, A.W., Moeller, A., Meyer, P., Sliz, P. and Richardson, C.C. (2017) Cryo-EM structure of the replisome reveals multiple interactions coordinating DNA synthesis. *Proc. Natl. Acad. Sci. U. S. A.*, **114**, E1848–E1856.
37. Wallen, J.R., Zhang, H., Weis, C., Cui, W., Foster, B.M., Ho, C.M.W., Hammel, M., Tainer, J.A., Gross, M.L. and Ellenberger, T. (2017) Hybrid methods reveal multiple flexibly linked DNA polymerases within the bacteriophage T7 replisome. *Structure*, **25**, 157–166.
38. Tanguy Le Gac, N., Villani, G., Hoffmann, J.S. and Boehmer, P.E. (1996) The UL8 subunit of the herpes simplex virus type-1 DNA helicase–primase optimizes utilization of DNA templates covered by the homologous single-strand DNA-binding protein ICP8. *J. Biol. Chem.*, **271**, 21645–21651.
39. Falkenberg, M., Bushnell, D.A., Elias, P. and Lehman, I.R. (1997) The UL8 subunit of the heterotrimeric herpes simplex virus type 1 helicase–primase is required for the unwinding of single strand DNA-binding protein (ICP8)-coated DNA substrates. *J. Biol. Chem.*, **272**, 22766–22770.
40. Ramanagoudr-Bhojappa, R., Blair, L.P., Tackett, A.J. and Raney, K.D. (2013) Physical and functional interaction between yeast Pif1 helicase and Rim1 single-stranded DNA binding protein. *Nucleic Acids Res.*, **41**, 1029–1046.
41. Nikali, K., Suomalainen, A., Saharinen, J., Kuokkanen, M., Spelbrink, J.N., Lonnqvist, T. and Peltonen, L. (2005) Infantile onset spinocerebellar ataxia is caused by recessive mutations in mitochondrial proteins Twinkle and Twinky. *Hum. Mol. Genet.*, **14**, 2981–2990.
42. Sarzi, E., Goffart, S., Serre, V., Chretien, D., Slama, A., Munnich, A., Spelbrink, J.N. and Rotig, A. (2007) Twinkle helicase (PEO1) gene mutation causes mitochondrial DNA depletion. *Ann. Neurol.*, **62**, 579–587.
43. Longley, M.J., Humble, M.M., Sharief, F.S. and Copeland, W.C. (2010) Disease variants of the human mitochondrial DNA helicase encoded by C10orf2 differentially alter protein stability, nucleotide hydrolysis, and helicase activity. *J. Biol. Chem.*, **285**, 29690–29702.
44. Morino, H., Pierce, S.B., Matsuda, Y., Walsh, T., Ohsawa, R., Newby, M., Hiraki-Kamon, K., Kuramochi, M., Lee, M.K., Klevit, R.E. *et al.* (2014) Mutations in Twinkle primase–helicase cause Perrault syndrome with neurologic features. *Neurology*, **83**, 2054–2061.
45. Tang, G., Peng, L., Baldwin, P.R., Mann, D.S., Jiang, W., Rees, I. and Ludtke, S.J. (2007) EMAN2: an extensible image processing suite for electron microscopy. *J. Struct. Biol.*, **157**, 38–46.
46. Marabini, R., Masegosa, I.M., San Martin, M.C., Marco, S., Fernandez, J.J., de la Fraga, L.G., Vaquerizo, C. and Carazo, J.M. (1996) Xmipp: an image processing package for electron microscopy. *J. Struct. Biol.*, **116**, 237–240.
47. Sorzano, C.O., Marabini, R., Velazquez-Muriel, J., Bilbao-Castro, J.R., Scheres, S.H., Carazo, J.M. and Pascual-Montano, A. (2004) XMIPP: a new generation of an open-source image processing package for electron microscopy. *J. Struct. Biol.*, **148**, 194–204.
48. Scheres, S.H., Gao, H., Valle, M., Herman, G.T., Eggermont, P.P., Frank, J. and Carazo, J.M. (2007) Disentangling conformational states of macromolecules in 3D-EM through likelihood optimization. *Nat. Methods*, **4**, 27–29.
49. Scheres, S.H., Nunez-Ramirez, R., Sorzano, C.O., Carazo, J.M. and Marabini, R. (2008) Image processing for electron microscopy single-particle analysis using XMIPP. *Nat. Protoc.*, **3**, 977–990.
50. Scheres, S.H., Valle, M., Grob, P., Nogales, E. and Carazo, J.M. (2009) Maximum likelihood refinement of electron microscopy data with normalization errors. *J. Struct. Biol.*, **166**, 234–240.