PNAS RESEARCH ARTICLE | BIOCHEMISTRY

Structural insight and characterization of human Twinkle helicase in mitochondrial disease

Amanda A. Riccio[a], Jonathan Bouvette[b], Lalith Perera[c], Matthew J. Longley[a], Juno M. Krahn[d], Jason G. Williams[e], Robert Dutcher[f], Mario J. Borgnia[b], and William C. Copeland[a,1]

Edited by Jodi Nunnari, University of California, Davis, Davis, CA; received April 29, 2022; accepted July 6, 2022

Twinkle is the mammalian helicase vital for replication and integrity of mitochondrial DNA. Over 90 Twinkle helicase disease variants have been linked to progressive external ophthalmoplegia and ataxia neuropathies among other mitochondrial diseases. Despite the biological and clinical importance, Twinkle represents the only remaining component of the human minimal mitochondrial replisome that has yet to be structurally characterized. Here, we present 3-dimensional structures of human Twinkle W315L. Employing cryo-electron microscopy (cryo-EM), we characterize the oligomeric assemblies of human full-length Twinkle W315L, define its multimeric interface, and map clinical variants associated with Twinkle in inherited mitochondrial disease. Cryo-EM, crosslinking-mass spectrometry, and molecular dynamics simulations provide insight into the dynamic movement and molecular consequences of the W315L clinical variant. Collectively, this ensemble of structures outlines a framework for studying Twinkle function in mitochondrial DNA replication and associated disease states.

mitochondrial DNA | Twinkle helicase | mitochondrial DNA replication | cryo-electron microscopy | progressive external ophthalmoplegia

Significance

Twinkle, the helicase in the mitochondrial DNA replisome, unwinds DNA in a 5′–3′ direction and improves the efficiency of the DNA polymerase and replication. To date, there are no atomic models of Twinkle and it is the last member of the mitochondrial DNA replisome to be structurally characterized. Given that mutations in Twinkle helicase cause clinical effects, mapping the location of these residues is critical to understanding Twinkle-dependent mitochondria disease phenotypes. Here, we utilize cryo-electron microscopy to determine a full-length structure of Twinkle W315L in two oligomeric states. In addition to defining the location of the clinical mutations, we investigate flexibility at the interface and propose a model for the clinical effects of W315L.

The human mitochondrial DNA (mtDNA) genome is 16.6 kb and encodes 13 proteins of the electron transport chain, 22 transfer RNAs, and 2 ribosomal RNAs. The human mitochondrial replisome includes: (1) the polymerase, Polγ, composed of the catalytic subunit, POLG, and its accessory subunit, POLG2, (2) Twinkle, (3) mitochondrial single-stranded binding protein, and (4) the mitochondria-directed RNA polymerase (POLRMT) (1, 2). Over 400 protein variants of this replisome have been shown to lead to mitochondrial dysfunction. Of these proteins, Twinkle helicase variants lead to infantile onset spinocerebellar ataxia and disease phenotypes characterized by neuromuscular dysfunction (3, 4). The essential function of Twinkle is also highlighted through its embryonic lethality in conditional knockout mice (5). Disease variants of Twinkle helicase can lead to increased mtDNA copy number (6), mtDNA depletion, and accumulation of replication intermediates (7).

Twinkle, a Superfamily 4 (SF4) helicase, catalyzes nucleoside 5′-triphosphate (NTP) hydrolysis and unwinds DNA substrates with 5′–3′ polarity (8, 9). Structurally, Twinkle is generally defined by: (1) an N-terminal region (NTR), (2) a primase-like domain, (3) a linker region, and (4) a C-terminal helicase domain (Fig. 1*A*). It is predicted that Twinkle shares overall fold similarity with other members of the SF4 helicases, such as T7 replicative helicase (gp4) with which it shares 46% sequence similarity and 15% identity. Most of the similarity is contained within the helicase domain (10). Similar to other SF4 helicases, structural models and biochemical data predict that Twinkle exists in various oligomeric states and undergoes oligomeric transitions [reviewed in (11)].

Despite sequence similarity in the C-terminus, the NTR and the primase-like domain of Twinkle lack critical residues to form zinc finger DNA binding domain and the residues central to primase function that are found within gp4 and other SF4 helicases (12). Thus, the N-terminal regions of Twinkle are poorly modeled (13). Although the quaternary assembly of Twinkle remains unclear, biochemical studies indicate that the NTR of Twinkle is susceptible to proteolysis and is, at least partially, responsible for DNA binding and oligomerization (12–15). The lack of high-resolution structural knowledge has hampered a thorough understanding of the molecular mechanisms of Twinkle helicase and Twinkle-dependent mtDNA deletions and mitochondrial disease.

Here, we present structures of human Twinkle helicase using the W315L disease variant. Obtained without crosslinking, the full-length (FL) human Twinkle W315L structures reveal distinct oligomeric states: heptamer and octamer. Both states of Twinkle W315L are refined and presented in cryo-electron microscopy (cryo-EM) maps with areas of high local resolution ~3.2 and 2.9 Å and ~4.5 and ~3.9 Å global map

Author contributions: A.A.R., M.J.L., and W.C.C. designed research; A.A.R., J.B., L.P., J.G.W., and R.D. performed research; A.A.R., J.B., L.P., M.J.L., J.M.K., J.G.W., R.D., M.J.B., and W.C.C. analyzed data; and A.A.R., J.B., L.P., M.J.L., J.M.K., J.G.W., R.D., M.J.B., and W.C.C. wrote the paper.

The authors declare no competing interest.

This article is a PNAS Direct Submission.

[1]To whom correspondence may be addressed. Email: copelan1@niehs.nih.gov.

This article contains supporting information online at http://www.pnas.org/lookup/suppl/doi:10.1073/pnas.2207459119/-/DCSupplemental.

Published August 1, 2022.

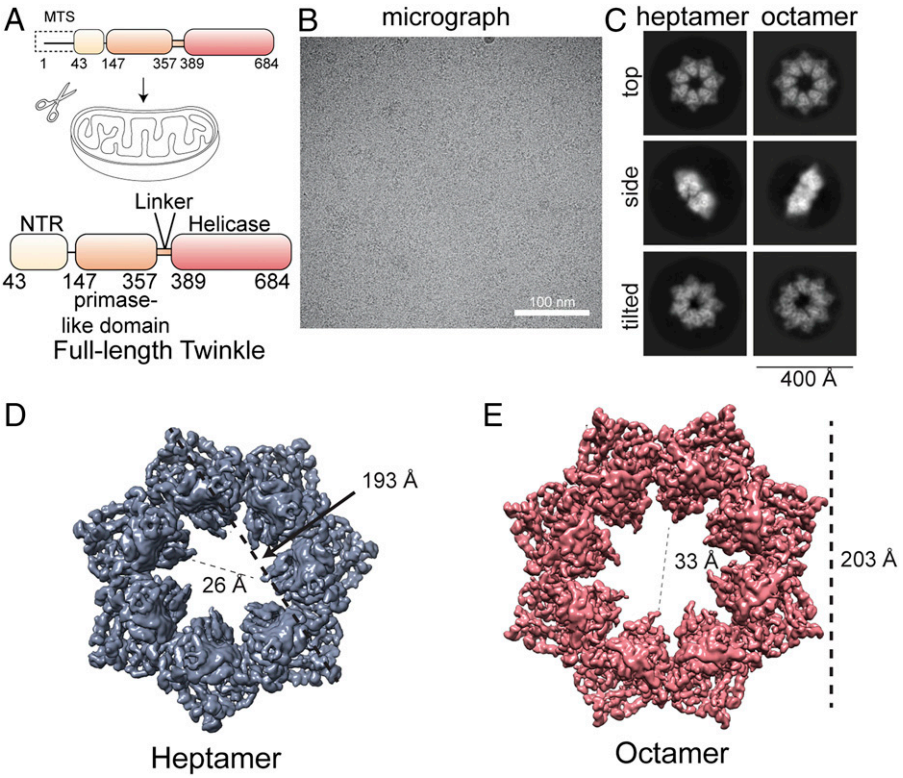

Fig. 1. Twinkle W315L analysis for structural studies. (*A*) Schematic of the Twinkle construct used in this study. Transport of Twinkle to the mitochondria removes the predicted mitochondrial transport sequence (MTS) producing Twinkle helicase (43-684), Twinkle FL ΔMTS). Twinkle domain structure consists of an N-terminal region (NTR) (light peach), a primase-like domain (light orange), and a C-terminal helicase domain (bright red). (*B*) Representative cryo-EM micrograph with scale bar (white). (*C*) Example cryo-EM 2D class averages of a heptamer and octamer Twinkle helicase. (*D* and *E*) Cryo-EM density maps of (*D*) heptamer (blue-gray) and (*E*) octamer (salmon) colored with size measurements indicated in angstroms of the lumen and width as measured in IMOD (49).

resolution for heptamers and octamers, respectively. These oligomeric structures are supported through crosslinking-mass spectrometry (XL-MS) and size-exclusion chromatography coupled to multi-angle light scattering (SEC-MALS). Our structures reveal a conformational assembly that is distinct from other known SF4 helicase conformations. Utilizing the cryo-EM-derived structures, molecular dynamics (MD) simulations define molecular details of wild-type (WT) Twinkle and the clinically identified W315L variant. Furthermore, our results provide spatial locations of disease variants and structural basis for dysfunctional mitochondrial DNA (mtDNA) replication observed in patients with Twinkle-associated mitochondrial diseases.

Results

Purification and Assembly of Multimeric Human Twinkle. Previous oligomeric assessments show WT and W315L eluting at high-molecular weight by size-exclusion chromatography (7). Both display a similar heptameric conformation upon negative stain EM, while WT also displays hexamers (4). Variable effects of W315L substitution on DNA helicase function *in vitro* have been reported and likely arise from differential assay sensitivities, reaction conditions, and DNA substrates (7, 10, 16). To aid in our understanding of autosomal dominant progressive external ophthalmoplegia (adPEO) resulting from mutations in c10orf2 (*TWNK* gene), we expressed full-length human W315L Twinkle (C-terminal his-tag Δ42, FL residues 43–684, 73 kDa) (16) in *Escherichia coli* (Fig. 1*A*), which we call Twinkle below. Detergents and glycerol, which have been among the common additives to aid WT and mutant Twinkle stability, have been removed (4, 12, 16). These additives often reduce the signal-to-noise ratio in micrographs, interfering with high-resolution cryo-EM (17). Our modified protocol (*Materials and Methods* and *SI Appendix*, Fig. 1*A*) permitted purification of Twinkle W315L in amounts and purity amenable to high-resolution structural analysis (*SI Appendix*, Fig. 1*B*).

Size-exclusion chromatography (SEC) coupled with multi-angle light scattering (SEC-MALS) of Twinkle in the absence (APO) and presence (DNA bound) of "FORK1" replication fork DNA (*SI Appendix*, Fig. 1*C*) confirmed self-assembly of high-molecular weight Twinkle W315L species, as previously observed (18). SEC-MALS profiles show the elution of the APO form at 13.66 mL with a derived molecular weight (MW) of 606 kDa that approximates an octameric species of Twinkle W315L. The DNA bound form had a small shift in peak elution to 13.38 mL with a derived MW of 503 kDa (*SI Appendix*, Fig. 1*C*). This reduced size approximates Twinkle species bound to FORK1 DNA following a monomer loss, which supports the hypothesis that Twinkle oligomers eject one monomer upon binding DNA (11).

High-Resolution Cryo-EM of Human Twinkle FL. A straightforward purification of Twinkle and confidence in oligomeric assembly allowed for the vitrification, sample imaging, subsequent cryo-EM data collection (*SI Appendix*, Fig. 2*A*), and single-particle analysis (Fig. 1 *B*–*E* and *SI Appendix*, Figs. 2–6 and Table 1) of Twinkle W315L. Twinkle W315L particles are discernible on micrographs (Fig. 1*B*), and data processing of two-dimensional (2D) classes (*SI Appendix*, Fig. 2*B*) indicate a mixture of oligomeric states distributed among heptamers (~30%), octamers (27%), oligomers containing DNA (23%), and broken rings of variable size (19%) (Fig. 1*C* and *SI Appendix*, Fig. 2*C*). Although grids were prepared with samples containing FORK1 DNA and

AMP-PNP, high orientation bias in the particle distribution of the DNA-bound species lead to low-resolution 3D reconstruction and precluded the generation of final 3D models. However, despite the oligomeric heterogeneity, possibly accentuated by the absence of crosslinking agents, cryo-EM clearly captures two distinct high-resolution APO oligomeric states, heptamer and octamer (Fig. 1 and *SI Appendix*, Figs. 2*C* and 3–6).

Following 2D and 3D classifications, the FSC (0.143) of the cryo-EM map of Twinkle W315L octamer and heptamer defined a global resolution ∼3.9 Å and ∼4.5 Å, respectively (C8, octamer EMD ID: 25743 and C7, heptamer EMD ID: 25744) (*SI Appendix*, Table 1 and Figs. 5 and 6). In these oligomeric assemblies, Twinkle W315L forms a ring-like structure (13, 14, 19). The heptamer has an outer diameter of ∼193 Å and an inner lumen diameter of ∼26 Å, and the octamer has an outer diameter of ∼203 Å and an inner lumen diameter of ∼33 Å (Fig. 1 *D* and *E*). The local resolution of the structure lies largely between ∼3.8–4.1 Å (heptamer) and ∼2.9–3.8 Å (octamer), with regions of presumably increased mobility at lower resolution, such as regions at the top and bottom of the ring (*SI Appendix*, Figs. 5 *B* and *C* and 6 *B* and *C*). Given the mobility of Twinkle in these regions and the lack of visual differences between individual monomers within the multimeric species, local refinement of the octamer was performed by masking a dimer unit and defining a fulcrum point at the center of the mask. The particles within each defining oligomeric structure were subject to symmetry expansion. With no applied symmetry (i.e., C1), the octameric cryo-EM map global resolution is ∼3.8 Å with local resolution ranging from ∼2.9 Å to 3.5 Å (*SI Appendix*, Fig. 3, EMD ID: 25746). The dimer map has improved resolution particularly in the regions forming the interior lumen of Twinkle.

Twinkle Ensemble. Given that there are no published structures of human Twinkle or homologs, we utilized the AlphaFold software and database (20) to generate models of both Twinkle W315L (AlphaFold software) and WT (AlphaFold database) as starting models for refinement (*SI Appendix*, Fig. 7). To investigate the structural details of Twinkle helicase W315L, we built and refined heptameric and octameric atomic models within the cryo-EM maps. We modeled seven asymmetric units into the heptameric cryo-EM map (PDB ID: 7T8C, EMD ID: 25744) (Fig. 2*A* and *SI Appendix*, Fig. 5*F*) and we modeled eight asymmetric units into the octameric cryo-EM density (PDB ID: 7T8B, EMD ID: 25743) (*SI Appendix*, Fig. 6*F*). The locally refined C1 map (EMD ID: 25746) was used to ensure model completeness. The density we observe for both the heptamer and octamer populations (*SI Appendix*, Figs. 5*F* and 6*F*, respectively) are well-fit by the Twinkle sequence. The Twinkle W315L structures are oligomeric ensembles of

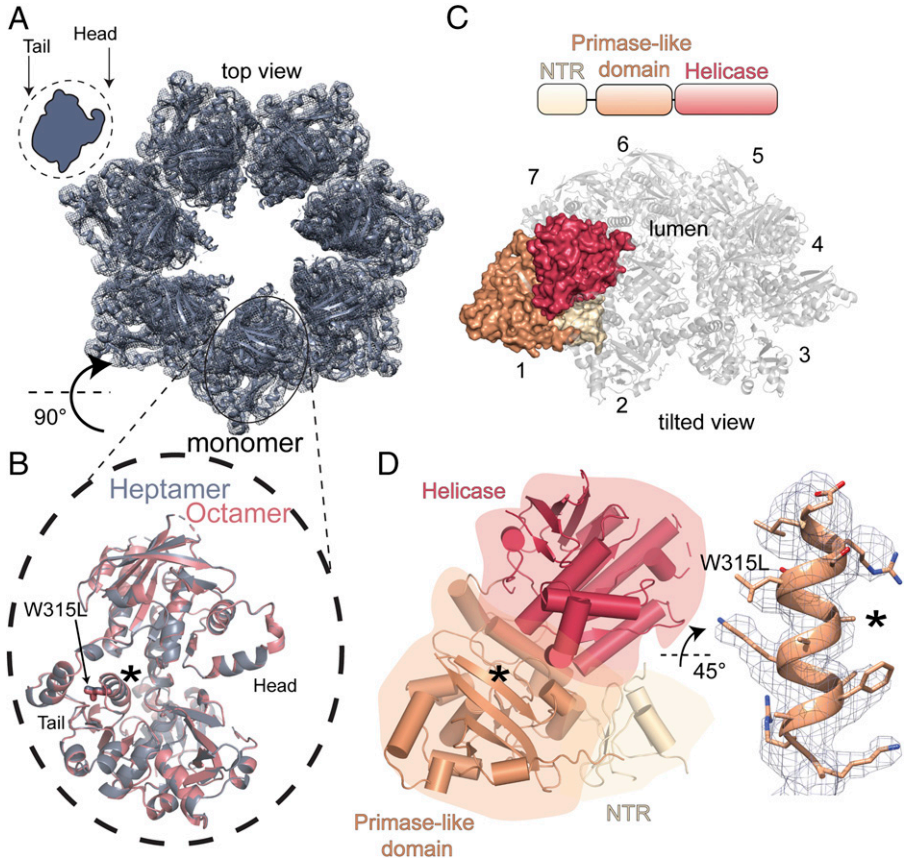

Fig. 2. Overview of W315L Twinkle helicase structure.(*A*) Top view of DeepEMhancer cryo-EM map of Twinkle helicase heptamer (EMD ID: 25744) (blue) fit with PDB ID: 7T8C (cartoon representation). The oval (*Top Left*) indicates an asymmetric unit or a monomer. (*B*) Zoom-in on a monomer unit of the heptamer PDB ID: 7T8C (blue) and octamer PDB ID: 7T8B (salmon) in cartoon representation. The asterisk indicates the helix containing W315L, and the arrow points to a stick representation of W315L. (*C*) (*Top*) Schematic of a monomer unit colored according to domain: N-terminal region (NTR) (light peach), primase-like domain (light orange), and C-terminal helicase domain (bright red). (*Bottom*) Mixed cartoon representation of a Twinkle heptamer (transparent gray) and a surface representation of a monomer (subdomain colors) within the heptamer. (*D*) (*Left*) Schematic of Twinkle subdomains mapped with cartoon representations of cylindrical helices. The cylindrical representation is colored as in (*C*), and the asterisk indicates the location of the helix containing W315L. (*Right*) The cryo-EM map (mesh) (EMD ID: 25743) fit with the PDB ID: 7T8B cartoon representation of the helix containing W315L. Residues are shown in stick representation. The asterisk is maintained in the same position in (*D*) and (*B*).

monomers containing "head" and "tail" regions that assemble in head-to-tail fashion to form a ring (Fig. 2 *A* and *B*). There is no observable density for residues 43–53, 85–98, 132–146, 227–230, and 633–684 or for the C-terminal his-tag. This absence of density is consistent with the disorder predicted for these regions by AlphaFold (20) (*SI Appendix*, Fig. 7) as well as DISOPRED, a disorder prediction software (21) (*SI Appendix*, Fig. 8).

The asymmetric units within the heptamer and octamer structures are nearly identical Twinkle monomers and chains A from each oligomer align with a root-mean-square deviation (RMSD) of 0.228 Å (Fig. 2*B*). Twinkle monomers within our Twinkle W315L ensembles are separated into three discrete subdomain regions: NTR (residues 54–131, light peach), primase-like domain and linker (residues 147–388, light orange), and C-terminal domain (residues 389–633, bright red) (Fig. 2 *C* and *D* and *SI Appendix*, Fig. 9). Additionally, the C-terminal domain stacks onto the N-terminal domain within each discrete monomer unit (Fig. 2 *C* and *D*). The Twinkle clinical variant W315L resides in the primase-like domain, and the helix containing W315L is well-occupied within the cryo-EM map (Fig. 2*D*). The positions of local residues surrounding W315L in the heptamer and octamer assemblies are remarkably similar (*SI Appendix*, Fig. 10), indicating that structural features are shared within the asymmetric units of both assemblies.

Crosslinking Mass Spectrometry of Twinkle Helicase. To validate our cryo-EM-derived model of Twinkle W315L, we performed chemical XL-MS. Under reducing SDS-PAGE conditions, bis[sulfosuccinimidyl]suberate (BS3d0) crosslinking yielded higher MW bands consistent with multimers, in addition to a ~73 kDa band for the noncrosslinked protein (Fig. 3*A*). In the samples analyzed by MS, there were 55 crosslinks and 27 unique crosslinks observed and well-defined (*SI Appendix*, Table 2 and Fig. 11). The Twinkle helicase sequence has an additional 19 lysine residues that were not observed in XL-MS data, highlighting the specificity of BS3d0 as a lysine-lysine crosslinker in this analysis. A XiNET plot displays all nonambiguous crosslinks (Fig. 3*B*) (22). The crosslinks were mapped onto the protein sequence and show clustering in three general regions. Using PDB ID: 7T8B chain B, the majority (88%) of crosslinks occurred within an acceptable region, defined by a cutoff of Cα-Cα ≤30 Å (Fig. 3*C*) (23–25). Within the permitted regions, crosslinks occurring near the multimerization interface are of particular interest. For example, crosslinks involving K319, K324, K328, and K365 (Fig. 3*D*) are all located within the predicted 12 Å region of BS3d0 and agree with the cryo-EM derived structures of Twinkle W315L. Each of the crosslinks occurring outside this cutoff is in connection to one residue, K593. Furthermore, the distances of crosslinking to this residue remain outside the cutoff in the context of both intra- and interdomain crosslinking. This residue is in the central lumen of the ring within the cryo-EM maps of Twinkle W315L, perhaps indicating flexibility in this region. A less populated conformation, not resolved to high-resolution in the cryo-EM analysis, could also explain the presence of this crosslink.

Molecular Dynamics Simulations of Conformational Flexibility, Stability, and Dynamic Changes. To understand the potential impact of the W315L clinical variant on the structure of Twinkle helicase, we performed molecular dynamics (MD) simulations using two simplified systems: (1) a dimer system (chain A and chain B), and (2) a monomer system (chain A) (Fig. 4*A*). We analyzed several parameters: RMSD, root-mean-square

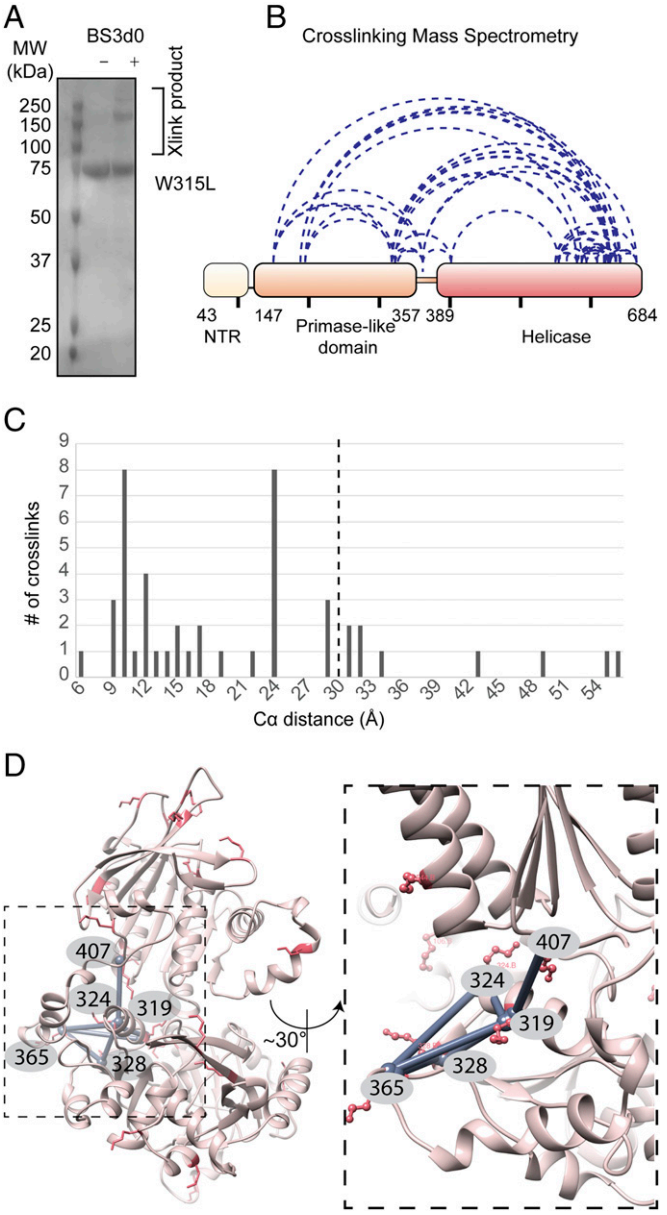

Fig. 3. XL-MS of Twinkle helicase W315L. (*A*) 4–12% SDS-PAGE with Coomassie staining of W315L Twinkle helicase in the presence and absence of Bs3d0 (0.5 mM). (*B*) 2D representation of Twinkle crosslinks displayed using XiNET (22). A schematic of the Twinkle residue numbers and domains are shown below the plot. The black tick marks indicate residues in increments of one hundred. (*C*) Histogram of linking distances derived from the map of Twinkle. Permissible crosslinks are defined as Cα-Cα ≤30 Å, distances were measured in Chimera. (*D*) (*Left*) A sample of Twinkle crosslinks modeled onto a Twinkle monomer (PDB ID: 7T8B). (*Right*) Crosslinked lysine side chains (sticks) near the multimerization interface are shown.

fluctuations (RMSF), and calculated dynamic cross correlation matrices (DCCM). The RMSD profile of Twinkle W315L as compared to WT is relatively unchanged within the dimer system (*SI Appendix*, Fig. 12). Similar deviations from the starting structures indicate the overall stability of both dimer systems is comparable. When treated as free monomers or monomer systems within a dimer, both systems are again stable (*SI Appendix*, Figs. 12 and 13*A*). When we truncate Twinkle in our system to remove the NTR (residues 56–149), we observe a slight decrease of RMSD values and therefore a mild increase in stability of the rest of the protein (compare *SI Appendix*, Figs. 14 vs. 12 and 13*A*). Meanwhile, the values of RMSDs calculated only from the NTR are increased for both WT and

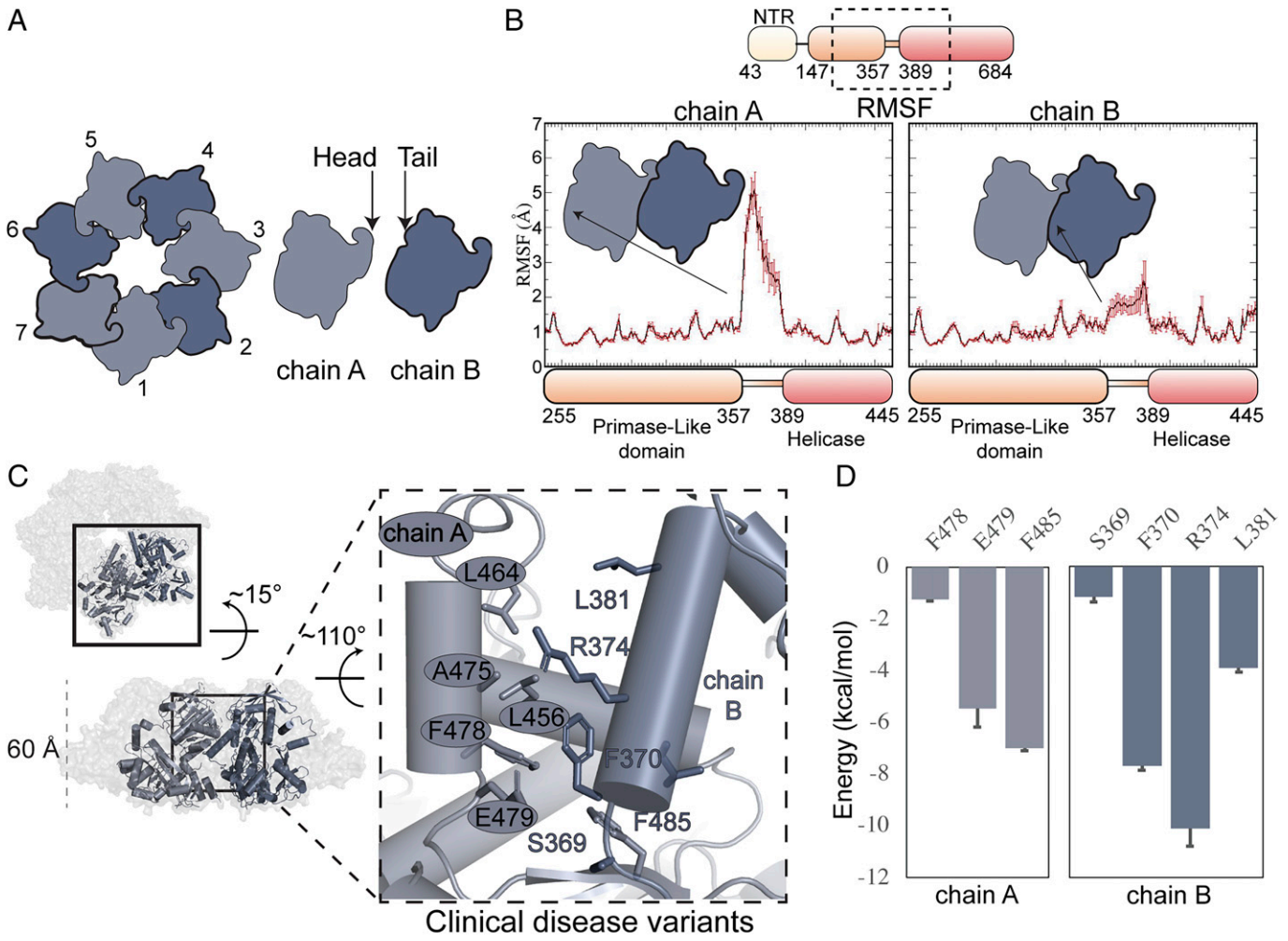

Fig. 4. Twinkle W315L oligomeric interface. Throughout the figure chain A of Twinkle is light gray blue and chain B is dark gray blue, in schematic a when displaying the oligomer chains c–g are colored alternating these same colors. (*A*) Schematic of Twinkle heptamer head-to-tail assembly and simplified dimer system used for MD simulations. (*B*) (*Top*) Schematic of Twinkle domain structure and the black dashed box pertaining to the region of the *Bottom Panel*. (*Bottom*) Zoom-in of the average RMSF plots of the Cα atoms for chain A (*Left*) and chain B (*Right*) of W315L calculated over the entire MD trajectory. X-axis is a schematic of the relevant domains of Twinkle and residue numbering. Arrows indicate the region on the schematic of an asymmetric unit pertaining to residues ~360–390 in the "tail region" of the Twinkle monomer for both chain A and chain B in a simplified dimer system. (*C*) A tilted view surface representation of Twinkle heptamer (PDB ID: 7T8C) and a black box highlights a cylindrical representation of two asymmetric units. The Twinkle heptamer is rotated to a side view of Twinkle, and a black box highlights the interface between asymmetric units. The asymmetric unit has a height of 60 Å as determined in Chimera. (*Right*) Zoom-in on the interface of chain A and chain B of PDB ID: 7T8C. Residues shown in stick representation are previously identified Twinkle clinical mutations (16). (*D*) A histogram of per-residue MM-GBSA binding energies of clinical variants at the interface of two asymmetric units of W315L. *y* axis represents the binding energies, and error bars are SE.

W315L in the monomer and dimer systems (*SI Appendix*, Fig. 15), indicating potential flexibility in the NTR.

We then explored the dynamic behavior of individual residues through analysis of RMSF of the α-carbons to predict local residue movement (Fig. 4*B* and *SI Appendix*, Fig. 13*B* and 16). There are negligible differences among the fluctuations when comparing chains A and B of both WT and W315L, with the exception of residues ~365 through 385, the "tail region" (Fig. 4 *A* and *B* and *SI Appendix*, Figs. 16 and 17 vs. 18, *Middle Panels*) and to a lesser extent residues ~460 through 495, the "head" region (*SI Appendix*, Figs. 16 and 17 vs. 18, *Lower Panels*). When the respective region is unengaged or unrestrained there is an elevated RMSF, conversely when it is restrained by the presence of a subsequent monomer, there is a decreased in RMSF. For example, when the tail (residues ~365–385) of chain A is unrestrained, there is a high RMSF (Fig. 4 *B*, *Left*), however, when the tail of chain B is bound to the head of chain A in an oligomer there is a reduction in RMSF at the interface (Fig. 4 *B*, *Right* and *SI Appendix*, Figs. 17 vs. 18, *Middle Panel*). Consistent with this observation, the monomer system has an elevated RMSF in both the head and tail region (*SI Appendix*, Figs. 13*B* and 19).

For additional insight into the dynamics of Twinkle WT and W315L, dynamic correlation was evaluated on the MD trajectories of WT and W315L (*SI Appendix*, Figs. 13*C*, 20, and 21). Dynamic correlation is reflective of the communication among various sub elements via α-carbons across the entire protein backbone. For Twinkle WT and W315L dimer and monomer systems, we clearly observe correlative movement indicative of three distinct subdomains within each chain, and residue numbering correlates well with the cryo-EM defined boundaries: NTR, primase-like domain, and the helicase domain (*SI Appendix*, Figs. 13*C*, 20, and 21). Here, WT exhibits marginally increased anti-correlative motions, or more flexibility as compared to W315L (*SI Appendix*, Figs. 20 vs. 21). Although there is anti-correlative movement, in the dimer systems, the local environment surrounding residue 315 in both W315 (WT) and W315L displays no major differences in local interaction frequencies in this

region (*SI Appendix*, Tables 3 and 4). From our MD simulations, we can see both WT and W315L systems are globally stable with regions of inherent flexibility.

Molecular Details of Twinkle Helicase Functioning. Guided by these structural features, molecular dynamics and cryo-EM become particularly advantageous to visualize specific residues and critical areas of Twinkle, such as the active site, oligomerization interfaces, and clinical mutations. Sequence and structural alignments of Twinkle and gp4 provide insight into the features of the active site, include the Walker A and B motifs (K421, E445, and D516), and the "arginine finger" R609 (*SI Appendix*, Fig. 22 *A* and *B*). While the NTP binding site of Twinkle is unoccupied, the residues necessary for nucleotide hydrolysis and NTP engagement are spatially assembled, and R609 is important for the interaction between subunits (*SI Appendix*, Fig. 22*C* and Table 5). Alignments of T7 helicase structure APO structure of the helicase-primase K318M mutant (PDB ID: 1Q57, gp4 residues 64–549), helicase E343Q catalytic mutant bound to single-strand DNA, and Mg^{2+}, dTTP (PDB ID: 6N7N, residues 264–549), Twinkle heptamer (PDB ID: 7T8C) and octamer (PDB ID: 7T8B, chains A) reveal similar active site assemblies (*SI Appendix*, Fig. 22 *D* and *E*).

These Twinkle W315L cryo-EM models also permit the mapping of an extensive interface between the "head" region (residues 448–459 and 464–479) and the "tail" region (residues 369–383) of the next monomer chain B (Fig. 4 *C* and *D*, heptamer). Comparing the dimer interfaces of the heptamer and octamer assemblies demonstrates that the interface is largely maintained and that global alignment of two asymmetric units has an RMSD of 0.288A (*SI Appendix*, Fig. 23*A*). More than 25 known missense mutations reside within the helices that form the subunit interface, and most of these variant Twinkle proteins have been shown in biochemical assays to interfere with multimerization or to reduce DNA helicase or ATPase activities (11, 16). Furthermore, at least 10 clinical missense variants directly align with the subunit interface in both the heptamer and the octamer (Fig. 4*C* and *SI Appendix*, Fig. 23*B*). Residues at the interface also display high frequencies of interaction (*SI Appendix*, Table 5 [chain B residues] and *SI Appendix*, Table 6 [chain A residues]). For example, the fractional occupancies of interdomain interactions for residues S369, F370, R374, L381 (chain B), F478, E479, and F485 (chain A) are between 0.921–1.000 for both WT and W315L Twinkle, indicating these residues are nearly always participating in the subunit interface. It is also apparent from molecular mechanics generalized Born surface area (MM/GBSA) or binding energy analysis that residues at the interface have large negative energies (*SI Appendix*, Fig. 23*C*, 24, and 25), including those implicated in clinical disease presentation (Fig. 4*D*). The binding energy analysis highlights the importance of these residues in assembly of Twinkle.

One of the most devastating clinical mutations of Twinkle is Y508C. Interestingly, Y508 is situated at the nexus of the N-terminal and C-terminal domain interface and it is also closely positioned to I367 (chain A) of the subsequent monomer (i.e., chain B), (*SI Appendix*, Fig. 26 *A*–*C*). The negative binding energies and high frequency of interaction of Y508 (chain A) (*SI Appendix*, Fig. 24 and Table 6) and I367 (chain B) (*SI Appendix*, Fig. 25 and Table 5) in MD simulations further support the notion that proper assembly of this region is critical for monomer formation, as well as oligomeric assembly. It is worth noting that in this region, I367T and V507I are also clinical variants associated with adPEO and Perrault syndrome 5, respectively. In total, the structures defined in this study also permitted the mapping of ~93% of disease-associated Twinkle mutations identified in Peter et al. (11) (*SI Appendix*, Fig. 26*D*).

Discussion

Architecture of Twinkle Helicase. We determined structures of the human mitochondrial replicative helicase, Twinkle W315L. The W315L disease variant structures illustrate the interface between Twinkle subunits and reveal contacts that are critical for the assembly of the oligomeric ring structure. Our cryo-EM structures display the primase-like domain stacking on the NTR within each monomer. Although our structures resemble previous 2D class analyses for Twinkle (4), their domain arrangement differs from a previously proposed Twinkle model (13). We did attempt to model a conformation of Twinkle in which the primase domain is rotated and positioned under the helicase domain of the adjacent monomer in the ring assembly, but this conformation did not fit appropriately into the cryo-EM map.

Previous studies have shown that the addition of NTPs and decreased ionic strength "compacts" Twinkle WT (4, 15). In contrast, W315L was not observed to have NTP-induced conformational changes. In these studies, 2D classes, under crosslinking conditions, display an NTR and/or primase-like domain folded inward toward the ring (4). In our study, the 2D classes and 3D structures of Twinkle W315L are consistent with a "compacted" conformation. While we cannot exclude that the presence of AMP-PNP in our sample is driving this compaction, we also do not observe a transition from open to extended conformations of Twinkle W315L, consistent with previous analysis. Additionally, Peter et al. (4) states that the "open" or star conformation of WT is unlikely to exist in vivo, but it likely implies an inherent flexibility that is absent in W315L. Our MD simulations systems indicate minor stability increases for W315L over WT (e.g., DCCM, *SI Appendix*, Figs. 20 and 21). If WT Twinkle has a propensity to sample an open conformation more frequently in vitro, this could contribute to the decreased solubility and increased aggregation observed by numerous researchers (11, 26). We suggest this inherent flexibility of WT Twinkle frustrated our previous structural efforts, whereas the relative stability of the W315L Twinkle enabled determination of structures with all domains visible in the oligomeric assemblies.

Oligomer Selection. Previous Twinkle models and descriptions of adPEO mutations have invoked altered flexibility and varied oligomeric states. Twinkle monomers, dimers, trimers, tetramers (82%) and larger oligomers (18%) have been observed by atomic force microscopy (19). Heterogenous mixtures of hexamer, heptamer, and broken rings were revealed by negative stain EM (4), and by cryo-EM, and hexamers have been observed after high salt GraFix and crosslinking (13). The environment and methods of protein purification may have generated these inconsistencies. Here, we have taken care to remove DNA from original purifications, to maintain salt conditions that closely resemble *in vivo* salt conditions, to eliminate detergents in sample buffers and to avoid biases associated with population selection or crosslinking. Perhaps the relative rigidity of the Twinkle W315L interface leads to a more conformationally homogenous population of Twinkle and promotes the appearance of an octamer population. The MD simulations data in this study, along with the cryo-EM and SEC-MALS data, suggest that flexibility at the

interface of the oligomeric species may allow for monomer exchange. Similarly, the flexibility observed at the interface may facilitate ring opening. In support of this, S369P and L381P substitutions at the subunit interface show increased broken rings (4). Our structures support biological observations that mutations to residues such as R374 (the highest binding energy to the dimer interface Fig. 4*D*) and Y508C would dramatically impact the ability of Twinkle to properly assemble.

Here, we compare atomic models of Twinkle in light of the well-characterized T7 gp4 structures, both in the absence and presence of the replisome components. Thus far, high-resolution structures of T7 gp4 have primarily analyzed the primase and helicase domains in the absence of the zinc finger binding domain (ZBD) [e.g., PDB IDs: 1Q57, primase-helicase; heptamer, (27); 1E0K, helicase (28); 5IKN, primase-helicase (29); and 1CR2, helicase (30)]. In the absence of the NTR, gp4 structures display primarily hexamers, although a heptamer was also observed in crystallization (27). Biochemical analysis suggests that the ZBD of T7 gp4 is inhibitory to hexamer formation and diffraction-competent crystals (29, 31).

It is interesting to postulate whether the NTR is inhibitory for Twinkle as well. If this is correct, then the removal of the Twinkle NTR might also produce a dramatic population shift toward a hexamer. Likewise, ZBD-primase domains of T7 gp4 have been resolved in structural studies [X-ray crystallography, PDB ID: 1NUI (32); ZBD-primase and cryo-EM, PDB ID: 6N9U (33); one ZBD, residues 9–45; and 2 primase domains residues (1) 64–209/218–260 and (2) 64–260], but the position of each ZBD within an oligomer assembly is largely unknown.

The Twinkle helicase conformation(s) in complex with the other members of the mitochondria replisome is, to our knowledge, unexplored. The T7 gp4 helicase-primase PDB ID: 5IKN vs. 1Q57 replisome structures describe nearly identical positioning in the presence and absence of additional replisome factors with an RMSD of 3.4 Å. It is also possible, as Peter et al. (11) proposed, that the C-terminus of Twinkle interacts with the other replisome components. Given that the C terminus is partially solvent exposed perhaps the current conformation of Twinkle is competent for protein-protein interaction. Experimental evidence of these interactions is under investigation.

Unrestrained Movement and Movement along DNA. Theories of helicase engagement and movement along DNA would support the need for sampling a broken ring conformation of Twinkle and sampling unrestrained movement observed in MD simulations. Similar unrestrained movement and open ring formation has been observed in the AFM study of Twinkle helicase (19) and high-resolution structural studies of T7 gp4 [e.g., PDB ID: 6N7I (33)]. Here, the observed change of MW calculated by SEC-MALS, in the presence of FORK1 DNA, is highly reminiscent of the ejection theory, whereby Twinkle helicase releases a monomeric subunit upon DNA binding. The clinical presentation of Twinkle W315L may arise from a dysfunction in DNA loading, whereby the oligomeric population would have subsequent rounds of monomer ejection, ejecting two molecules (APO octamer to DNA-bound hexamer) (Fig. 5). Additional structural studies will be needed to characterize the mechanism of Twinkle helicase movement along DNA, in the context of adPEO clinical variants and WT. While it is possible that the diameter of the lumen observed in our Twinkle structures could accommodate both the single-strand and double-strand elements of the DNA fork, we do not observe unmodeled density in the ring-like 3D reconstructions that could account for the presence of DNA. Under the present conditions, the Twinkle-DNA bound species in our cryo-EM dataset suggest that the structure underwent a large-scale conformational change, but did not yield a high-resolution map due to orientation bias in the projection distribution (*SI Appendix*, Fig. 2*C*). Previously, work has shown that Twinkle helicase is capable of binding DNA in the absence of cofactors (12, 34) and here, the sample reaction assembly was designed similarly to Patel and colleagues (35), in which helicase activity would begin with the addition of Mg^{2+}. Future studies are needed to further explore the conformational changes and biochemical properties in cryo-EM of Twinkle in the presence and absence of cofactors, metals, and nucleic acids.

Summary. Overall, this study unveils high-resolution structures of Twinkle helicase as both a heptamer and an octamer. We are now able to accurately map nearly all clinically identified Twinkle variants. XL-MS experiments and MD simulations further support the cryo-EM structure. How Twinkle helicase assembles within the mitochondrial replisome, identification of protein-to-protein or protein-to-DNA contacts, and how Twinkle dysfunction promotes replisome-dependent stalling and ultimately the formation of mtDNA deletions are ongoing investigations. We anticipate this study will elicit subsequent investigations of the conformational changes identifying additional states of Twinkle helicase, as well as assembly within the replisome. Furthermore, this study provides a platform for targeted therapies in the

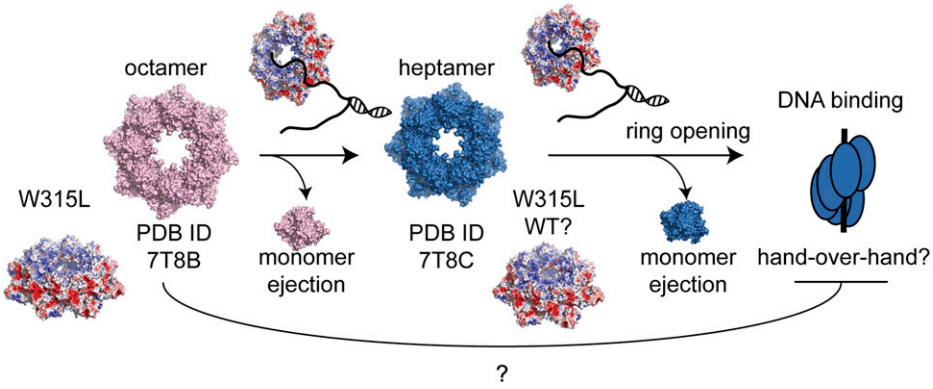

Fig. 5. Proposed model of Twinkle WT and W315L octamer-heptamer monomer ejection. The loss of one (WT/W315L) or two monomer (W315L) units is required for DNA binding. Each attempt to bind DNA allows for ring opening and unrestrained movement exploiting the inherent flexibility of Twinkle. Twinkle top views are shown in cartoon surface representation for the heptamer (blue) and octamer (pink). Tilted views are shown using the Adaptive Poisson-Boltzmann Solver (APBS) electrostatics plugin of PyMOL onto the surface representation of both oligomeric assemblies. 3D surface cartoon representation color-coded electrostatic surface −5 (red) to +5 (blue) K_bT/e_c.

treatment of Twinkle-associated diseases such as progressive external ophthalmoplegia, Perrault syndrome, infantile-onset spinocerebellar ataxia/IOSCA, and hepatocerebral mtDNA depletion syndrome.

Materials and Methods

Plasmids and Expression. Twinkle W315L plasmid pT684Δ42C-His, previously described (16), was transformed in Rosetta2 (DE3) (Millipore). Cultures were grown in terrific broth to an OD_{600} of ∼2 and induced with 0.2 mM IPTG. Cultures were incubated overnight at 16 °C and shaking at 180 rpm. Of note, these analyses described here were also attempted for Twinkle WT, but, as is well-established, solubility prevented further analysis. Therefore, for simplicity, W315L is referred to as Twinkle.

W315L Purification. Bacterial pellet was resuspended in cold lysis buffer (25 mL/L culture) (25 mM Hepes pH 8, 500 mM NaCl, 5% glycerol, 0.5 mM TCEP, 0.5% Nonidet P-40, 40 mM imidazole) and cells were disrupted on ice by sonication until homogenous. The crude cell extract was clarified by centrifugation at 20 k × *g*, 4 °C and applied to a fresh 5 mL Ni-NTA HP column (Cytiva). After washing the column with lysis buffer and lysis buffer without Nonidet P-40, protein was eluted in 5 mL fractions of no Nonidet P-40 lysis buffer supplemented with 250 mM imidazole. The purest fractions were pooled and diluted to 250 mM NaCl and loaded on a fresh 1 mL heparin HP column (Cytiva). The protein was eluted in a linear gradient 15 CV Buffer A (25 mM Hepes pH 8, 250 mM NaCl, 0.5 mM TCEP) to Buffer B (Buffer A supplemented with 1 M NaCl) in 2-mL fractions. For cryo-EM, the peak fraction was dialyzed at 4 °C for 1 h in Buffer A. For SEC-MALS, fractions containing the purest protein were supplemented with 50 mM imidazole and 50 mM NaCl, concentrated to 5 mL using Amicon filter 50 K cutoff (Millipore) and injected on a SEC S200 16/60 hi-prep column (Cytiva) in Twinkle imidazole buffer (25 mM Hepes pH 8, 300 mM NaCl, 50 mM imidazole, 0.5 mM TCEP). All proteins used in biochemical assays were flash frozen and stored at −80 °C. Proteins for SEC-MALS were "polished" over an S200 column in Twinkle imidazole buffer prior to SEC-MALS experiments. Cryo-EM grids were prepared from freshly purified protein.

DNA Substrate Complex Formation. Fork1 substrate was incubated at 95 °C for 5 min and annealed by slowly decreased to room temperature in 1× annealing buffer (10 mM Hepes pH 8, 50 mM NaCl, 10 mM EDTA). DNA fork substrate was previously described (18) and contains the following oligonucleotides: (1) dscrandom44: ACTTGAATGCGGCTTAGTATGCATTGTAAAACGACGGCCAGTGC, (2) TSTEM GCACTGGCCGTCGTTTTACGGTCGTGACTGGGAAAACCCTGGCG, and (3) U25 CGCCAGGGTTTTCCCAGTCACGACC (Integrated DNA Technologies).

Cryo-EM Freezing and Sample and Grid Preparation. A 2-nm carbon film was evaporated on G-flat R1.2/1.3 (Protochips, Inc) using a Leica ACE600. Grids were glow-discharged on both sides for 30 s, 15 mA using a Pelco EasiGlow at 0.38 mBar. 3 μL of buffer (220 mM NaCl, 25 mM Hepes pH 8, 0.5 mM TCEP) was applied to the back side of the grid. The Twinkle reaction (3 μL) containing freshly purified Twinkle W315L ΔMTS (43-684) (∼1.2–1.8 mg/mL) with 1 mM β,γ-imidoadenosine 5′-triphophate lithium salt hydrate (Sigma Aldrich), 20 μM Fork1 DNA diluted with 250 mM NaCl, and 25 mM Hepes pH 8, 0.5 mM TCEP which was preincubated for 1 h on ice was applied to the front side. Excess liquid was blotted for 3–4 s, blotting force of −1, 95% humidity, and 12 °C using a Vitrobot Mark II (Thermo Fisher) and plunged into liquid ethane.

Cryo-EM Data Collection, Processing, and Refinement. Datasets were collected on a Talos Arctica microscope (Thermo Fisher) at 200 kV using a K2 camera (Gatan Inc.) at a pixel size of 0.932 Å/px and on a Titan Krios microscope (Thermo Fisher) at 300 kV equipped with a bio-quantum energy filter (Gatan Inc) and either a K2 or a K3 detector (Gatan Inc) at a super resolution pixel size of 0.529 Å/px at 0 and 30° tilt angle (*SI Appendix*, Fig. 2*A*). Movies from all datasets were aligned using MotionCor2 (36) using 5 × 5 tiles and binned to a common pixel size of 1.058 Å/px. Aligned micrographs were imported and processed into CryoSPARC v3.0 (37). Final maps were sharpened using DeepEMhancer (38). Atomic models were built into the maps with Chimera (39), Coot (40), Phenix (41), and PyMOL (42).

Chemical Crosslinking and Mass Spectrometry. Twinkle W315L (7 μM, monomer) was incubated with 1X annealing buffer or FORK1 DNA for 15 min. Crosslinking reactions started upon the addition of BS3d0 to a final concentration of 0.5 mM. Reaction mixtures were incubated for 5 min at room temperature. Reactions were quenched with the addition of Tris pH 8 (20 mM final) followed by 15 min on ice. For sample to be run on SDS-PAGE gel, Laemmli buffer was added prior to loading on to 4–12% SDS-PAGE gels (Invitrogen). Mass Spectrometry analysis was performed as described in reference (43). 10 μL of the crosslinked complex was digested by addition of 10 μL trypsin/LysC mix (0.11 μg/μL - Promega) and held overnight at 37 °C. Digests were stored at -80 °C for subsequent MS analysis. Protein digests were analyzed by LC/MS on a Q Exactive Plus mass spectrometer (Thermo Fisher Scientific) interfaced with an M-Class nanoAcquity UPLC system (Waters Corporation) equipped with a 75 μm × 150 mm HSS T3 C18 column (1.8 μm particle, Waters Corporation) and a C18 trapping column (180 μm × 20 mm) with 5-μm particle size at a flow rate of 450 nL/min. The trapping column was in-line with the analytical column and upstream of a microtee union which was used for venting, trapping, and as a liquid junction. Trapping was performed using the initial solvent composition. 5 μL of digested sample was injected onto the column. Peptides were eluted by using a linear gradient from 99% solvent A (0.1% formic acid in water [vol/vol]) and 1% solvent B (0.1% formic acid in acetonitrile [vol/vol]) to 40% solvent B over 70 min. For the mass spectrometry, a top–ten data-dependent acquisition method was employed with a dynamic exclusion time of 15 s and exclusion of singly charged species. The mass spectrometer was employed with a nanoflex source with a stainless-steel needle and used in the positive ion mode. Instrument parameters were as follows: sheath gas, 0; auxiliary gas, 0; sweep gas, 0; spray voltage, 2.7 kV; capillary temperature, 275 °C; S-lens, 60; scan range (*m/z*) of 375–1,500; 1.6 *m/z* isolation window; resolution: 70,000 (MS), 17,500 (MS/MS); automated gain control (AGC), $3 \times 10e^6$ ions (MS), $5 \times 10e^4$ (MS/MS); and a maximum IT of 100 ms (MS), 50 ms (MS/MS). Mass calibration was performed before data acquisition using the Pierce LTQ Velos Positive Ion Calibration mixture (Thermo Fisher Scientific). The LC/MS raw data were first converted to an MGF format using Mascot Distiller from Matrix Science and then analyzed using the Batch-Tag Web function of the ProteinProspector web-based software developed by the UCSF Mass Spectrometry Facility. The MGF file was searched against sequences for the recombinant Twinkle Helicase by employing the User Protein Sequence field with other search parameters including: tryptic specificity and three missed cleavages; precursor charge range of 2, 3, 4, and 5; monoisotopic values; parent mass tolerance of 20 ppm and fragment mass tolerance of 50 ppm; oxidation of methionine as a variable modification; and in the Crosslinking field, the Link Search Type was defined as DSS. The putative crosslinked peptide output was triaged by limiting the mass error of putative crosslinks to two SDs from the average error (∼5 ppm); requiring a Score Difference value >4 except for the cases of intermolecular crosslinks of identical peptides or peptides ≤3 amino acid residues; and total expectation values below 1×10^{-4}.

Molecular Dynamics Simulations. The initial model of the WT system was constructed by mutating L315 to W using Coot. For the preparation of starting configurations of molecular dynamics trajectories, which were based on PDB ID: 7T8B, missing protons were introduced by using the leap module of Amber.18 (44), six Na^+ counter ions were added to the dimer system, and the systems were solvated in a box of TIP3P water (45) with the box boundary extending to 20 Å from the nearest protein atom (resulting in a total of 186,579 atoms in the rectangular simulation box). Based on their local environment (by manually examining the potential for making hydrogen bonds with the neighboring residues), histidine residues 67, 405, 503, 512, and 552 were selected to be δ-protonated while the rest of the histidine residues were ε-protonated. There were 102 Na^+ and 102 Cl^- ions added to the solution providing the 100 mM effective salt concentration closer to the physiological conditions. Prior to equilibration, the solvated system was sequentially subjected to (1) 500 ps of dynamics of water and all ions with fixed (or frozen) protein, (2) 5,000 steps of energy minimization of all atoms (the first 2,000 steps with the steepest descent method followed by the conjugate gradient method for the rest), (3) an initial constant temperature (300 K) — constant pressure (at 1 atm with the isotropic position scaling) dynamics at fixed protein to assure a reasonable starting density (∼1 ns) while keeping the protein positions under constraints with a

10 kcal/mol force constant for 50 ps, (4) a conjugate-gradient minimization for 1,000 steps, (5) step-wise heating (from 0 to 300 K in 3 ns) MD at constant volume, and (6) constant volume simulation for 10 ns with a constraint force constant of 10 kcal/mol applied only on backbone heavy atoms. After releasing all constraining forces within the next 40 ns equilibration period, sampling was increased by performing 5 independent molecular dynamics simulations for 1.20 μs at constant temperature (300 K) and constant volume. The constant temperature was maintained using the Langevin dynamics with the collision frequency of 0.5 ps^{-1}. A similar protocol was adopted for the five 1.20 μs trajectories of both the Twinkle WT and W315L. The box dimensions were around 137.49 Å × 111.23 Å × 124.15 Å. For monomer simulations, three Na^+ counterions were added. There were additional 84 Na^+ ions and 84 Cl^- ions providing an effective salt concentration of 100 mM. Total number of atoms of the WT-monomer system was 147,576. Five 1.20 μs independent simulations were performed for each of the WT and the W315L monomer systems. All trajectories were calculated using the PMEMD module of Amber.18 with 1-fs time step. Long-range coulombic interactions were handled using the PME method (46) with the cutoff of 10 Å for the direct interactions. The amino acid parameters were selected from the FF14SB force field (45, 47) of Amber.18. RMSD, RMSF, dynamic cross correlation matrices, and other analysis were performed using the CPPTRAJ module (48) of Amber.18 in conjunction with some in-house programs. At the salt concentration of 100 mM with the standard parameters, the MM/GBSA module, with the standard parameters, was used to estimate residue-residue interaction energies for all protein residues. A total of 6,000 configurations selected at each nanosecond from the five 1.2 μs MD trajectories for both the WT and W315L mutant dimer systems. All residues were used in the MMGBSA energy decomposition.

Data Availability. The cryo-EM maps and atomic coordinates have been deposited in the Electron Microscopy Data Bank (EMD ID: 25743, 25744, 25746) (50–52) and Protein Data Bank (PDB ID: 7T8B and 7T8C) (53,54). Materials, such as gels or molecular simulations, are available upon request to the corresponding author. All study data are included in the article and/or *SI Appendix*.

ACKNOWLEDGMENTS. We thank Drs. Margaret Gustafson and Venkata Dandey for their critical reading of this manuscript and insightful discussions. This work was supported by the Division of Intramural Research of the NIH, National Institute of Environmental Health Sciences (NIEHS) (Z01 ES065078 and Z01 ES065080 to W.C.C., Z01 ES043010 to L.P., and ZIC ES 103326 to M.J.B.). Molecular graphics and analyses performed with UCSF Chimera, developed by the Resource for Biocomputing, Visualization, and Informatics at the University of California, San Francisco, with support from NIH P41-GM103311.

Author affiliations: [a]Mitochondrial DNA Replication group, Genome Integrity and Structural Biology Laboratory, National Institute of Environmental Health Sciences, NIH, Research Triangle Park, NC 27709; [b]Molecular Microscopy Consortium, Genome Integrity and Structural Biology Laboratory, National Institute of Environmental Health Sciences, NIH, Research Triangle Park, NC 27709; [c]Computational Chemistry and Molecular Modeling Support Group, National Institute of Environmental Health Sciences, NIH, Research Triangle Park, NC 27709; [d]Genome Integrity and Structural Biology Laboratory, National Institute of Environmental Health Sciences, NIH, Research Triangle Park, NC 27709; [e]Mass Spectrometry Research and Support Group, Epigenetics and Stem Cell Biology Laboratory, Genome Integrity and Structural Biology Laboratory, National Institute of Environmental Health Sciences, NIH, Research Triangle Park, NC 27709; and [f]Macromolecular Structure Group, Epigenetics and Stem Cell Biology Laboratory, National Institute of Environmental Health Sciences, NIH, Research Triangle Park, NC 27709

1. J. A. Korhonen, X. H. Pham, M. Pellegrini, M. Falkenberg, Reconstitution of a minimal mtDNA replisome in vitro. *EMBO J.* **23**, 2423–2429 (2004).
2. S. Wanrooij *et al.*, Human mitochondrial RNA polymerase primes lagging-strand DNA synthesis in vitro. *Proc. Natl. Acad. Sci. U.S.A.* **105**, 11122–11127 (2008).
3. A. H. Hakonen *et al.*, Recessive Twinkle mutations in early onset encephalopathy with mtDNA depletion. *Brain* **130**, 3032–3040 (2007).
4. B. Peter, G. Farge, C. Pardo-Hernandez, S. Tångefjord, M. Falkenberg, Structural basis for adPEO-causing mutations in the mitochondrial TWINKLE helicase. *Hum. Mol. Genet.* **28**, 1090–1099 (2019).
5. D. Milenkovic *et al.*, TWINKLE is an essential mitochondrial helicase required for synthesis of nascent D-loop strands and complete mtDNA replication. *Hum. Mol. Genet.* **22**, 1983–1993 (2013).
6. H. Tyynismaa *et al.*, Twinkle helicase is essential for mtDNA maintenance and regulates mtDNA copy number. *Hum. Mol. Genet.* **13**, 3219–3227 (2004).
7. S. Goffart *et al.*, 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 (2009).
8. J. A. Korhonen, M. Gaspari, M. Falkenberg, TWINKLE Has 5′ -> 3′ DNA helicase activity and is specifically stimulated by mitochondrial single-stranded DNA-binding protein. *J. Biol. Chem.* **278**, 48627–48632 (2003).
9. J. N. Spelbrink *et al.*, 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 (2001).
10. T. Holmlund *et al.*, Structure-function defects of the twinkle amino-terminal region in progressive external ophthalmoplegia. *Biochim. Biophys. Acta* **1792**, 132–139 (2009).
11. B. Peter, M. Falkenberg, TWINKLE and other human mitochondrial DNA helicases: Structure, function and disease. *Genes (Basel)* **11**, 408 (2020).
12. G. Farge *et al.*, The N-terminal domain of TWINKLE contributes to single-stranded DNA binding and DNA helicase activities. *Nucleic Acids Res.* **36**, 393–403 (2008).
13. P. Fernández-Millán *et al.*, The hexameric structure of the human mitochondrial replicative helicase Twinkle. *Nucleic Acids Res.* **43**, 4284–4295 (2015).
14. L. S. Kaguni, M. T. Oliveira, Structure, function and evolution of the animal mitochondrial replicative DNA helicase. *Crit. Rev. Biochem. Mol. Biol.* **51**, 53–64 (2016).
15. T. D. Ziebarth *et al.*, Dynamic effects of cofactors and DNA on the oligomeric state of human mitochondrial DNA helicase. *J. Biol. Chem.* **285**, 14639–14647 (2010).
16. M. J. Longley, M. M. Humble, F. S. Sharief, W. C. Copeland, 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 (2010).
17. I. Drulyte *et al.*, Approaches to altering particle distributions in cryo-electron microscopy sample preparation. *Acta Crystallogr. D Struct. Biol.* **74**, 560–571 (2018).
18. I. Khan *et al.*, Biochemical characterization of the human mitochondrial replicative Twinkle helicase: Substrate specificity, DNA branch migration, and ability to overcome blockades to DNA unwinding. *J. Biol. Chem.* **291**, 14324–14339 (2016).
19. P. Kaur *et al.*, Single-molecule level structural dynamics of DNA unwinding by human mitochondrial Twinkle helicase. *J. Biol. Chem.* **295**, 5564–5576 (2020).
20. A. W. Senior *et al.*, Improved protein structure prediction using potentials from deep learning. *Nature* **577**, 706–710 (2020).
21. J. J. Ward, J. S. Sodhi, L. J. McGuffin, B. F. Buxton, D. T. Jones, Prediction and functional analysis of native disorder in proteins from the three kingdoms of life. *J. Mol. Biol.* **337**, 635–645 (2004).
22. C. W. Combe, L. Fischer, J. Rappsilber, xiNET: Cross-link network maps with residue resolution. *Mol. Cell. Proteomics* **14**, 1137–1147 (2015).
23. E. D. Merkley *et al.*, Distance restraints from crosslinking mass spectrometry: Mining a molecular dynamics simulation database to evaluate lysine-lysine distances. *Protein Sci.* **23**, 747–759 (2014).
24. S. M. Vos *et al.*, Structure of activated transcription complex Pol II-DSIF-PAF-SPT6. *Nature* **560**, 607–612 (2018).
25. M. Jones *et al.*, Cryo-EM structures of the XPF-ERCC1 endonuclease reveal how DNA-junction engagement disrupts an auto-inhibited conformation. *Nat. Commun.* **11**, 1120 (2020).
26. T. D. Ziebarth, C. L. Farr, L. S. Kaguni, Modular architecture of the hexameric human mitochondrial DNA helicase. *J. Mol. Biol.* **367**, 1382–1391 (2007).
27. E. A. Toth, Y. Li, M. R. Sawaya, Y. Cheng, T. Ellenberger, The crystal structure of the bifunctional primase-helicase of bacteriophage T7. *Mol. Cell* **12**, 1113–1123 (2003).
28. M. R. Singleton, M. R. Sawaya, T. Ellenberger, D. B. Wigley, Crystal structure of T7 gene 4 ring helicase indicates a mechanism for sequential hydrolysis of nucleotides. *Cell* **101**, 589–600 (2000).
29. J. R. Wallen *et al.*, Hybrid methods reveal multiple flexibly linked DNA polymerases within the bacteriophage T7 replisome. *Structure* **25**, 157–166 (2017).
30. M. R. Sawaya, S. Guo, S. Tabor, C. C. Richardson, T. Ellenberger, Crystal structure of the helicase domain from the replicative helicase-primase of bacteriophage T7. *Cell* **99**, 167–177 (1999).
31. H. Zhang, S.-J. Lee, A. W. Kulczyk, B. Zhu, C. C. Richardson, Heterohexamer of 56- and 63-kDa gene 4 helicase-primase of bacteriophage T7 in DNA replication. *J. Biol. Chem.* **287**, 34273–34287 (2012).
32. M. Kato, T. Ito, G. Wagner, C. C. Richardson, T. Ellenberger, Modular architecture of the bacteriophage T7 primase couples RNA primer synthesis to DNA synthesis. *Mol. Cell* **11**, 1349–1360 (2003).
33. Y. Gao *et al.*, Structures and operating principles of the replisome. *Science* **363**, eaav7003 (2019).
34. D. Sen, D. Nandakumar, G.-Q. Tang, S. S. Patel, Human mitochondrial DNA helicase TWINKLE is both an unwinding and annealing helicase. *J. Biol. Chem.* **287**, 14545–14556 (2012).
35. G. Patel *et al.*, A257T linker region mutant of T7 helicase-primase protein is defective in DNA loading and rescued by T7 DNA polymerase. *J. Biol. Chem.* **286**, 20490–20499 (2011).
36. S. Q. Zheng *et al.*, MotionCor2: Anisotropic correction of beam-induced motion for improved cryo-electron microscopy. *Nat. Methods* **14**, 331–332 (2017).
37. A. Punjani, J. L. Rubinstein, D. J. Fleet, M. A. Brubaker, cryoSPARC: Algorithms for rapid unsupervised cryo-EM structure determination. *Nat. Methods* **14**, 290–296 (2017).
38. R. Sanchez-Garcia *et al.*, DeepEMhancer: A deep learning solution for cryo-EM volume post-processing. *Commun. Biol.* **4**, 874 (2021).
39. E. F. Pettersen *et al.*, UCSF Chimera–a visualization system for exploratory research and analysis. *J. Comput. Chem.* **25**, 1605–1612 (2004).
40. P. Emsley, K. Cowtan, Coot: Model-building tools for molecular graphics. *Acta Crystallogr. D Biol. Crystallogr.* **60**, 2126–2132 (2004).
41. D. Liebschner *et al.*, Macromolecular structure determination using X-rays, neutrons and electrons: Recent developments in Phenix. *Acta Crystallogr. D Struct. Biol.* **75**, 861–877 (2019).
42. The PyMOL Molecular Graphics System. Version 1.8. Schrödinger, LLC (2015).
43. M. C. Pillon *et al.*, Cryo-EM reveals active site coordination within a multienzyme pre-rRNA processing complex. *Nat. Struct. Mol. Biol.* **26**, 830–839 (2019).
44. D. A. Case, I. Y. Ben-Shalom, S. R. Brozell, D. S. Cerutti, T. E. Cheatham, III, V. W. D. Cruzeiro, T. A. Darden, R. E. Duke, D. Ghoreishi, M. K. Gilson, H. Gohlke, A. W. Goetz, D. Greene, R. Harris, N. Homeyer, Y. Huang, S. Izadi, A. Kovalenko, T. Kurtzman, T. S. Lee, S. LeGrand, P. Li, C. Lin, J. Liu, T. Luchko, R. Luo, D. J. Mermelstein, K. M. Merz, Y. Miao, G. Monard, C. Nguyen, H. Nguyen, I. Omelyan, A. Onufriev, F. Pan, R. Qi, D. R. Roe, A. Roitberg, C. Sagui, S. Schott-Verdugo, J. Shen, C. L. Simmerling, J. Smith, R. Salomon-Ferrer, J. Swails, R. C. Walker, J. Wang, H. Wei, R. M. Wolf, X. Wu, L. Xiao, D. M. York and P. A. Kollman (2018), AMBER 2018, University of California, San Francisco.

45. W. L. Jorgensen, J. Chandrasekhar, J. D. Madura, R. W. Impey, M. L. Klein, Comparison of simple potential functions for simulating liquid water. *J. Chem. Phys.* **79**, 926–935 (1983).
46. U. Essmann *et al.*, A smooth particle mesh Ewald method. *J. Chem. Phys.* **103**, 8577–8593 (1995).
47. W. D. Cornell *et al.*, A second generation force field for the simulation of proteins, nucleic acids, and organic molecules. *J. Am. Chem. Soc.* **117**, 5179–5197 (1995).
48. D. R. Roe, T. E. Cheatham III, PTRAJ and CPPTRAJ: Software for processing and analysis of molecular dynamics trajectory data. *J. Chem. Theory Comput.* **9**, 3084–3095 (2013).
49. J. R. Kremer, D. N. Mastronarde, J. R. McIntosh, Computer visualization of three-dimensional image data using IMOD. *J. Struct. Biol.* **116**, 71–76 (1996).
50. A. A. Riccio, J. Bouvette, J. M. Krahn, M. J. Borgnia, W. C. Copeland, Heptameric Human Twinkle Helicase Clinical Variant W315L. EMDB. https://www.ebi.ac.uk/pdbe/entry/emdb/EMD-25743. Deposited 16 December 2021.
51. A. A. Riccio, J. Bouvette, J. M. Krahn, M. J. Borgnia, W. C. Copeland, Octameric Human Twinkle Helicase Clinical Variant W315L. EMDB. https://www.ebi.ac.uk/pdbe/entry/emdb/EMD-25744. Deposited 16 December 2021.
52. A. A. Riccio, J. Bouvette, M. J. Borgnia, W. C. Copeland, Octameric Twinkle Helicase Clinical Variant W315L, local refinement. EMDB. https://www.ebi.ac.uk/pdbe/entry/emdb/EMD-25746. Deposited 17 December 2021.
53. A. A. Riccio, J. Bouvette, J. M. Krahn, M. J. Borgnia, W. C. Copeland, Heptameric Human Twinkle Helicase Clinical Variant W315L. RCSB PDB. PDB DOI: https://doi.org/10.2210/pdb7T8B/pdb. Deposited 16 December 2021.
54. A. A. Riccio, J. Bouvette, J. M. Krahn, M. J. Borgnia, W. C. Copeland, Octameric Human Twinkle Helicase Clinical Variant W315L. RCSB PDB. PDB DOI: https://doi.org/10.2210/pdb7T8C/pdb. Deposited 16 December 2021.