



Modular Architecture of the Hexameric Human Mitochondrial DNA Helicase

Tawn D. Ziebarth, Carol L. Farr and Laurie S. Kaguni*

Department of Biochemistry
and Molecular Biology
Michigan State University
East Lansing, MI 48823, USA

We have probed the structure of the human mitochondrial DNA helicase, an enzyme that uses the energy of nucleotide hydrolysis to unwind duplex DNA during mitochondrial DNA replication. This novel helicase shares substantial amino acid sequence and functional similarities with the bacteriophage T7 primase-helicase. We show in velocity sedimentation and gel filtration analyses that the mitochondrial DNA helicase exists as a hexamer. Limited proteolysis by trypsin results in the production of several stable fragments, and N-terminal sequencing reveals distinct N and C-terminal polypeptides that represent minimal structural domains. Physical analysis of the proteolytic products defines the region required to maintain oligomeric structure to reside within amino acid residues ~405–590. Truncations of the N and C termini affect differentially DNA-dependent ATPase activity, and whereas a C-terminal domain polypeptide is functional, an N-terminal domain polypeptide lacks ATPase activity. Sequence similarity and secondary structural alignments combined with biochemical data suggest that amino acid residue R609 serves as the putative arginine finger that is essential for ATPase activity in ring helicases. The hexameric conformation and modular architecture revealed in our study document that the mitochondrial DNA helicase and bacteriophage T7 primase-helicase share physical features. Our findings place the mitochondrial DNA helicase firmly in the DnaB-like family of replicative DNA helicases.

© 2007 Elsevier Ltd. All rights reserved.

Keywords: mitochondria; DNA replication; DNA helicase; bacteriophage T7; hexamer

*Corresponding author

Introduction

A nuclear gene known as *Twinkle* encodes the human mitochondrial DNA (mtDNA) helicase.¹ A mitochondrial link to human disease first allowed its identification in individuals afflicted with autosomal dominant progressive external ophthalmoplegia, a disease of external eye muscle paralysis associated with multiple mtDNA deletions. That *Twinkle* might be the mitochondrial replicative helicase was suggested by its apparent amino acid (aa) sequence homology with bacteriophage T7 gene 4 protein (T7

gp4), a bi-functional primase-helicase.² Biochemical characterization of a recombinant human mtDNA helicase demonstrated its NTPase and DNA unwinding capabilities.³ UTP was found to be the preferred nucleotide substrate, and a 5'-stretch of single-stranded DNA (ssDNA) and a short 3'-ssDNA tail are required for translocation and duplex unwinding in a 5'-3' direction. Physiological studies reveal that the enzyme is essential for mtDNA maintenance, and regulation of mtDNA copy number.^{1,2}

All helicases contain a conserved alpha-beta domain, known as the RecA-like core. This core region houses the functional Walker A and Walker B motifs that contain the residues responsible for nucleotide hydrolysis.^{4,5} Furthermore, these enzymes comprise additional distinct domains illustrating an overall modular architecture. The accessory domains contribute to mechanistic features by providing directionality or substrate specificity, and include the broad functions of the helicases, which are all members of the larger family of AAA⁺ ATPases.⁴

Abbreviations used: mtDNA, mitochondrial DNA; aa, amino acid; T7 gp4, bacteriophage T7 gene 4 protein; ssDNA, single-stranded DNA; ZBD, zinc binding domain; RPD, RNA polymerase domain; dsDNA, double-stranded DNA; STI, soybean trypsin inhibitor; PMSF, phenylmethylsulfonyl fluoride.

E-mail address of the corresponding author:
lskaguni@msu.edu

The T7 gp4 is a well-characterized, bifunctional enzyme that serves as a model for studies of novel replicative helicases. It functions as a heterohexamer with protomers that comprise five conserved helicase motifs, placing it in the *Escherichia coli* DnaB-like family of replicative helicases.^{4,6} T7 gp4 shares this association with other hexameric helicases such as bacteriophage T4 gene 41 protein (T4 gp41) and the RepA protein encoded by plasmid RSF1010.^{7–10} Electron microscopic studies of these multimers reveal a subunit ring assembly that produces a central channel for the threading of DNA.^{6,11,12}

Early physical studies of T7 gp4 and *E. coli* DnaB protein identified common proteolytically-sensitive regions.^{13–15} The full-length T7 gp4 comprises three distinct structural and functional domains that are separated by flexible linkers. Its N-terminal domain contains two subdomains, a zinc binding domain (ZBD) and an RNA polymerase domain (RPD).^{16,17} The ZBD recognizes and binds to a DNA priming site (5'-GTC-3') and subsequently delivers the primed-template to the DNA polymerase active site.¹⁷ The RPD synthesizes an oligoribonucleotide primer that is extended by the T7 DNA polymerase, T7 gene 5 protein.^{18,19} The C-terminal helicase domain of T7 gp4 contains the dTTPase activity and is solely responsible for unwinding double-stranded DNA (dsDNA) by displacing the complementary strand.^{13,19,20} The N and C-terminal domains are separated by a linker that is important for oligomerization, and also mediates conformational changes.^{21,22}

E. coli DnaB protein functions as a DNA helicase during bacterial DNA replication and interacts with its corresponding primase, *E. coli* DnaG protein.^{9,23} The DnaB protein monomer contains two structural domains separated by a hinge region.¹⁴ Its N-terminal domain dimerizes under certain conditions to regulate the quaternary structural changes that the protein undergoes, and is essential for helicase activity.^{12,24,25} The hinge participates in protein-protein interactions with DnaG protein, and residues in the C-terminal domain comprise its DNA binding, oligomerization and ATPase activities.^{1,14,24}

Variations among the DnaB-like enzymes include the regions responsible for oligomerization. The majority of subunit interactions in T7 gp4 involve helix A (aa 272–281) of one monomer, which lies C-terminal to the linker region, interacting with a neighboring monomer *via* helices D1–D3 (aa 345–388), which lie in a region between motif IA and motif II. Additional interaction contributions come from specific residues in the linker region.^{21,22,26} In the RSF1010-encoded RepA protein, the extreme N terminus comprising residues 2–18 of one monomer interacts with residues 108–114 of the adjacent monomer.¹¹ In contrast, the N termini of both T7 gp4 and DnaB protein are readily cleaved by limited proteolysis with retention of oligomeric structure.^{13,14}

High-resolution crystal structures are available for several members of the DnaB-like family of replicative helicases, including T7 gp4 and RepA protein.^{11,27} The T7 gp4 lacking its N-terminal ZBD

shows a structural asymmetry that suggests a mechanism of sequential hydrolysis to explain its translocation along DNA.²⁰ RepA protein, on the other hand, shows complete symmetry among its subunits.¹¹ Findings similar to those for RepA protein were reported in electron microscopic studies of the *E. coli* DnaB protein under specific conditions.¹² The differences in subunit orientation among the helicases argue against a common mechanism of nucleotide hydrolysis leading to translocation.^{4,26,28,29} Instead, such mechanistic diversity may be relevant for cellular versatility as demonstrated in other hexameric ATPases.^{30–32}

To evaluate the structural features of the human mtDNA helicase, we pursued a combined approach of limited proteolysis and hydrodynamic methods. We demonstrate the modular architecture of the hexameric enzyme. We evaluate our findings in comparison with other functionally homologous proteins, placing the human mtDNA helicase in the DnaB-like family of replicative helicases.

Results

The human mtDNA helicase has a hexameric quaternary structure

The human mtDNA helicase lacking its mitochondrial targeting sequence (aa 43–684) was purified previously to near-homogeneity in four chromatographic steps from baculovirus-infected insect cells.³ We developed a streamlined purification strategy using N and C-terminally hexa-histidine-tagged forms of the protein for our physical studies (see Materials and Methods). This procedure yielded 0.7 mg of near-homogeneous enzyme per liter of infected cells.

To investigate the oligomeric state of our recombinant human mtDNA helicase, we used hydrodynamic methods to determine its molecular mass. Velocity sedimentation produced a single peak with a sedimentation coefficient of 13.6 S (Figure 1(a)). Similarly, gel filtration of the enzyme produced a single peak with a Stokes radius of 80 Å (Figure 1(b)). Combining the sedimentation coefficient with the Stokes radius yields a native molecular mass of 420,000 daltons. SDS-PAGE of peak fractions from velocity sedimentation and gel filtration reveals a 72,000 dalton polypeptide. Taken together, our data indicate that the purified, recombinant human mtDNA helicase exists as a hexamer, and is consistent with the prominent oligomeric state of the members of the *E. coli* DnaB-like family.

The hexameric human mtDNA helicase is a modular protein

A modular architecture is a structural characteristic of T7 gp4 and *E. coli* DnaB protein.^{14,15} We investigated whether the human mtDNA helicase comprises

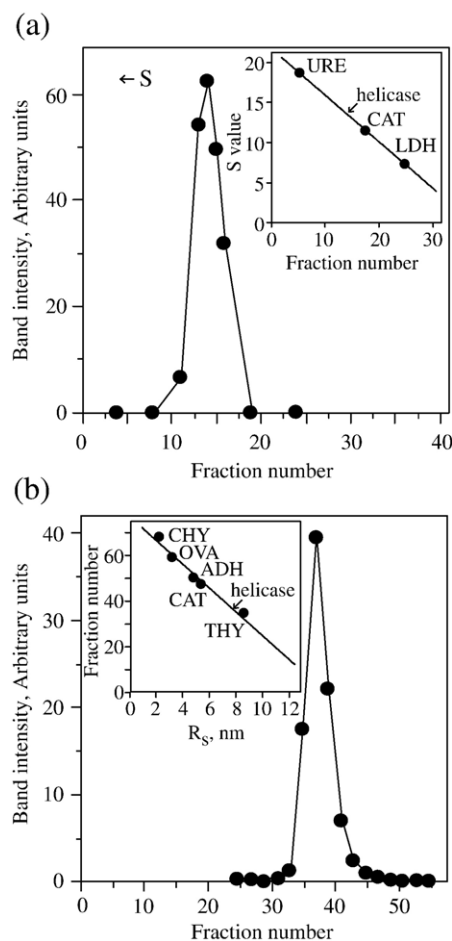


Figure 1. Hydrodynamic analysis of human mtDNA helicase. (a) Glycerol gradient sedimentation. Helicase was sedimented in a 12–30% glycerol gradient (see Materials and Methods). Standard protein markers run in parallel gradients were: jack bean urease (URE, 18.6 S), bovine liver catalase (CAT, 11.3 S), and rabbit muscle lactate dehydrogenase (LDH, 7.3 S). (b) Gel filtration. Helicase was chromatographed on a Superdex 200 column. Standard protein markers used to calibrate the column were: chymotrypsinogen (CHY, R_s 2.9 nm), hen egg ovalbumin (OVA, R_s 3.05 nm), yeast alcohol dehydrogenase (ADH, R_s 4.6 nm), bovine liver catalase (CAT, R_s 5.2 nm), and bovine thyroid thyroglobulin (THY, R_s 8.5 nm).

distinct structural domains by proteolyzing the enzyme using a titration of trypsin under limiting conditions, and identified six prominent cleavage products of 66 (T66), 57 (T57), 50 (T50), 34 (T34), 30 (T30), and 28 (T28) kDa (Figure 2(a)). Furthermore, limited proteolysis with a variety of proteases with diverse cleavage specificities showed that each of the products of trypsin digestion was produced by more than one protease (Figure 2(b)).

The initial trypsin product, T66, was evident upon silver staining and immunoblot analyses using anti-helicase antibody of both N and C-terminally His-tagged human mtDNA helicase. However, immunoblots probed with anti-His monoclonal antibody detected only the N-terminally His-tagged T66, illustrating cleavage and a loss of the His-tag from the

carboxyl terminus (data not shown). Based on these data, we conclude that T66 derives from cleavage in an arginine and lysine-rich region near aa residue 640, and represents a human mtDNA helicase lacking its C terminus (Figure 3(a)). For each of the proteolytic fragments, the position of the C terminus was estimated from the migration in SDS-PAGE, the cleavage specificity of trypsin and a secondary structural prediction indicating non-structured regions (Figure 3(b)). N-terminal sequencing demonstrates that T57 is the product of subsequent cleavage at the N terminus of T66. It is produced by cleavage at aa residue K144, and creates a polypeptide lacking both termini. Furthermore, N-terminal sequencing of the T50, T34 and T28 fragments demonstrates that they all share a common N terminus with T57. We conclude that there is a progressive digestion of T57 at its C terminus until a minimal fragment is produced, T28. T30 is produced from cleavage at aa residue R371, and represents a polypeptide spanning aa ~372–640. A time-course of trypsin digestion demonstrates the stable nature of each fragment, and shows that the two polypeptides T30 and T28 represent minimal N and C-terminal structural domains (Figure 4).

Oligomeric properties of the human mtDNA helicase proteolytic products

To elucidate the regions responsible for maintenance of the hexameric form of the enzyme, we

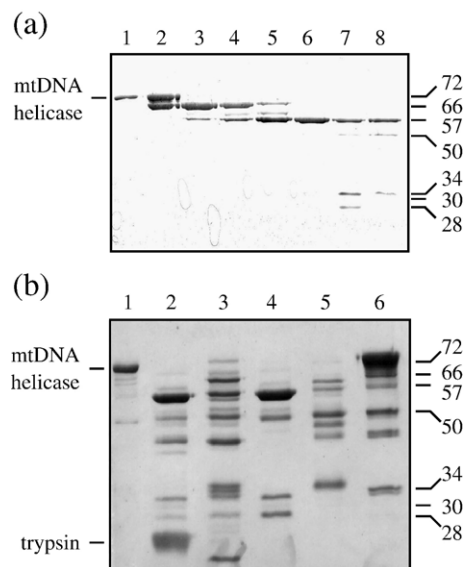


Figure 2. Proteolytic digestion produces stable fragments of human mtDNA helicase. (a) Trypsin titration. Helicase was digested with increasing concentrations of trypsin. Samples were analyzed by SDS-PAGE followed by silver staining. Ratios are given as helicase:trypsin. Lane 1, undigested helicase; lanes 2–8, 1: 0.02, 1: 0.08, 1: 0.2, 1: 0.3, 1: 0.6, 1: 1.3, 1: 2.0. (b) Multiple protease digestion. Helicase was digested with five different proteases, and the products were analyzed by SDS-PAGE followed by silver staining. Ratios are given as helicase:protease. Lane 1, undigested helicase; lane 2, trypsin 1:1; lane 3, chymotrypsin 1:0.35; lane 4, thermolysin 1:1; lane 5, subtilisin 1:0.005; lane 6, papain 1:0.18.

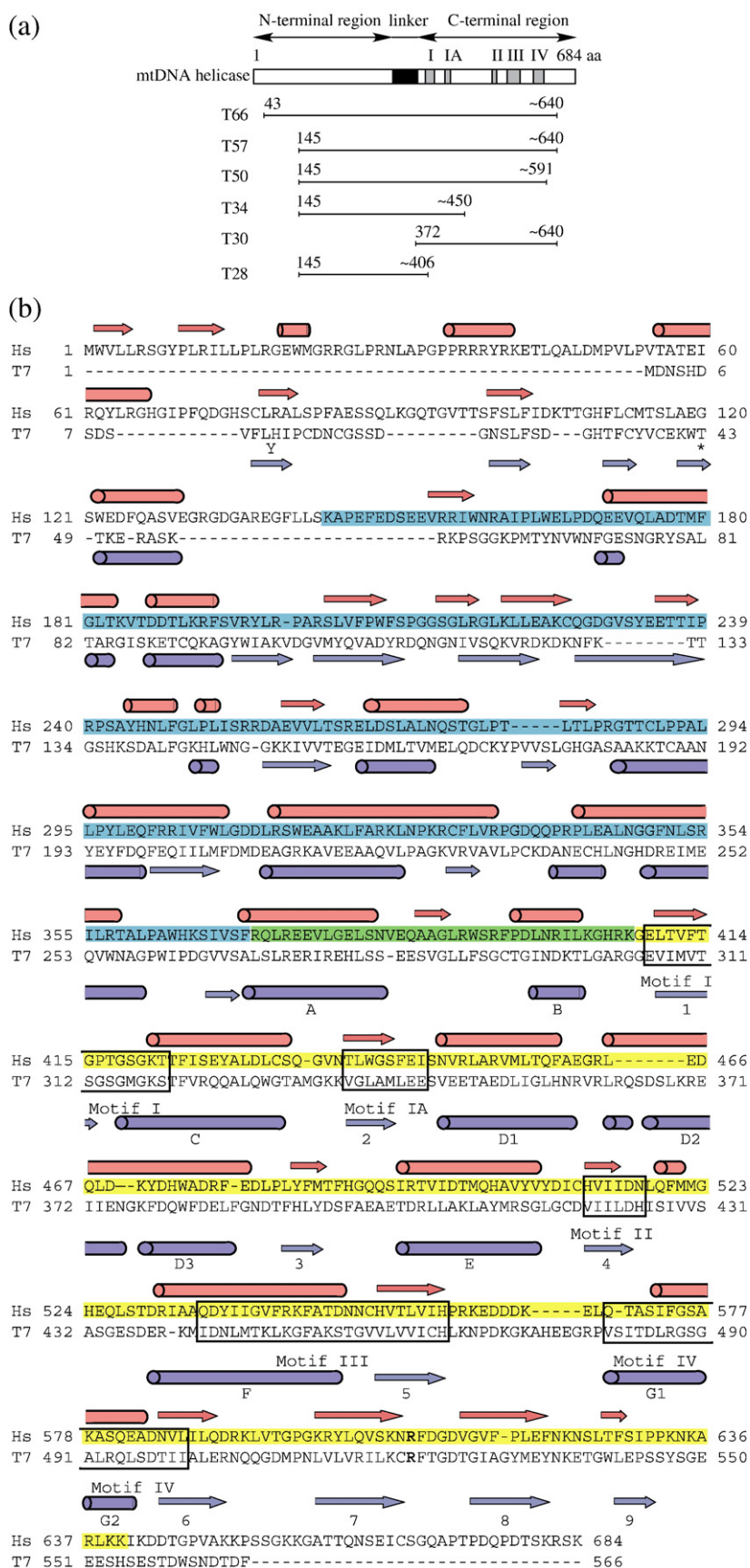


Figure 3. (a) Schematic map of proteolytic fragments of human mtDNA helicase. The recombinant form of the full-length human mtDNA helicase used here lacks its mitochondrial targeting sequence, and comprises amino acid residues 43–684. (b) Sequence alignment of human mtDNA helicase and bacteriophage T7 primase-helicase. A secondary structure-based sequence alignment of T7 gp4 (PDB codes 1NUI¹⁷ and 1Q57²⁷) and human mtDNA helicase (GenBank accession number NP068602) was produced using 3D-PSSM (www.sbg.bio.ic.ac.uk/~3dpssm). In the T7 gp4 amino acid sequence, residue Y13 is displaced below the alignment and amino acid residues 44–48 (AGNED) are omitted. Barrels designate alpha helices and arrows designate beta sheets. The five conserved amino acid motifs in superfamily IV DNA helicases are indicated by labeled boxes. The N-terminal T28 fragment of human mtDNA helicase is shaded in blue, the C-terminal T30 fragment is shaded in yellow and the overlapping region is in green. The position of the arginine finger in the helicase domain of T7 gp4 (R522) and the corresponding amino acid residue in the human mtDNA helicase (R609) are indicated in bold.

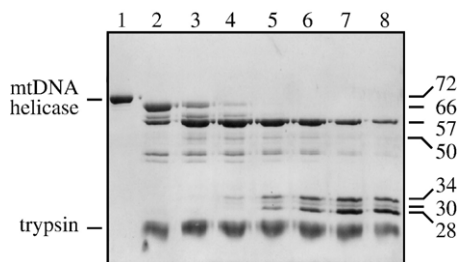


Figure 4. Time-course of trypsin digestion. Helicase was digested over a time-course of 1.5 h. Samples were extracted at the indicated time points, and analyzed by SDS-PAGE followed by silver staining. Lane 1, undigested helicase; lanes 2–8, 1, 5, 10, 20, 30, 60, and 90 min.

examined the oligomeric properties of mixtures of the proteolytic products, T66/T57, T57/T50 and T34/T28. Samples were subjected to gel filtration under the elution conditions used to obtain the Stokes radius of the intact enzyme. T66/T57 demonstrated an elution profile similar to the full-length enzyme with a peak in fractions 38–42, and no protein apparent in any other fraction (Figure 5(a)).

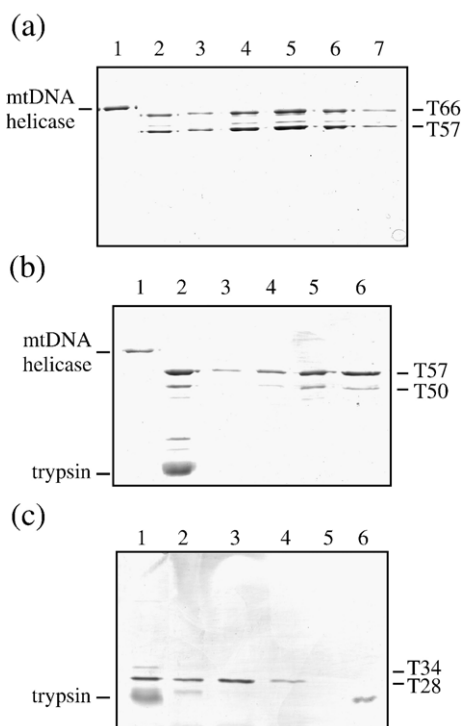


Figure 5. Gel filtration of trypsin products T66/T57, T57/T50 and T34/T28. Helicase was subjected to limited proteolysis with trypsin and the samples were fractionated by gel filtration on a FPLC Superdex 200 column. The elution fractions were analyzed by SDS-PAGE and silver staining; all those containing staining material are shown here. (a) T66/T57. Lane 1, undigested helicase; lane 2, gel filtration load; lanes 3–7, hexameric peak fractions 38–42. (b) T57/T50. Lane 1, undigested helicase; lane 2, gel filtration load; lanes 3–6, hexameric peak fractions 39–42. (c) T34/T28. Lane 1, gel filtration load; lanes 2–4, monomeric peak fractions 65–67; lane 5, blank; lane 6, trypsin.

Using an appropriate trypsin:helicase ratio to ensure complete proteolysis of T66 but optimal production of T57 and T50, we observed a single chromatographic peak in fractions 39–42, and again found no protein apparent in any other fraction (Figure 5(b)). SDS-PAGE and silver staining revealed both T57 and T50, consistent with their presence in a hexameric hetero-oligomer. We also used a trypsin:helicase ratio appropriate to ensure maximal production of T28 with minimal retention of the larger proteolytic products. Upon gel filtration of this digest we observed a single chromatographic peak eluting in fractions 65–67, consistent with the presence of T28 eluting as a monomer (Figure 5(c)). Therefore, we conclude that the region required for retention of an oligomeric structure resides within aa ~405–590.

ATPase activity of the human mtDNA helicase and its proteolytic products

To reveal structure-function relationships in the proteolyzed forms of the human mtDNA helicase, we evaluated their DNA-dependent ATPase activities relative to the full-length enzyme (Figure 6). First, we investigated whether the termini contribute to nucleotide hydrolysis by purifying recombinant proteins corresponding to the proteolytic products T66 and T57, which we refer to as P66 and P57, respectively (see Materials and Methods). Consistent with our initial gel filtration studies of the proteolytic products, we found the P66 and P57 proteins to be hexamers upon purification (data not shown). Remarkably, the homogeneous P66 that lacks its C terminus displays reproducibly increased ATPase activity as compared to the similar activities of full-length and the homogeneous P57, a polypeptide lacking both termini, in two preparations of each protein. These data suggest that the C terminus of the intact polypeptide is inhibitory to ATPase activity, whereas the N terminus may contribute to the overall catalytic activity of the human mtDNA

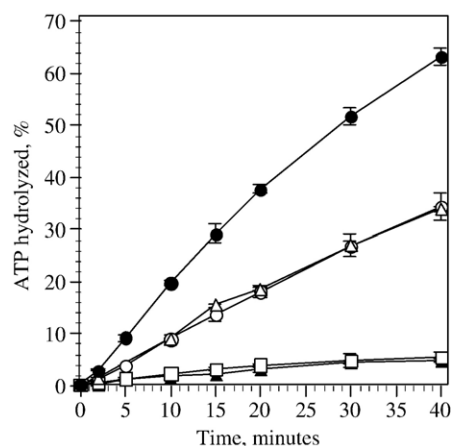


Figure 6. Time-course of ATP hydrolysis by trypsin products. Full-length helicase (open circles), P66 (filled circles), P57 (open triangles), T57/T50 (4:1, filled triangles and 3:1, open squares) were assayed for ATPase activity over 0–40 min.

helicase. In contrast, the heterogeneous oligomer of T57/T50, at ratios of 4:1 and 3:1, exhibits an ~fourfold decrease in activity. Taken together with the data obtained for the homogeneous P57, we conclude that the lower activity of the T57/T50 complex is due to the lack of critical aa residues in T50.

Discussion

The human mtDNA helicase shares significant sequence similarity and functional characteristics with T7 gp4.^{2,3} Multiple sequence alignments have demonstrated the conservation of the five highly conserved helicase motifs, and our secondary structural prediction of the mitochondrial enzyme aligned with the known structure of the bacteriophage enzyme using 3D-PSSM suggests additional sequence similarities that likely pertain to structure. We investigated this possibility by examination of the native structure of the recombinant human mtDNA helicase using velocity sedimentation and gel filtration. Spelbrink *et al.* showed previously that the human mtDNA helicase forms oligomers *in vivo*.² Our data demonstrate that the recombinant human mtDNA helicase exists as a hexamer under the ionic conditions of our purification, a common conformation for the active form of replicative helicases.^{7,33}

Our limited proteolysis study demonstrates that the human mtDNA helicase and bacteriophage T7 gp4 share similar structural domains. Remarkably, comparison of our trypsin titration of the human mtDNA helicase with an endoprotease Glu-C titration of T7 gp4 reveals indistinguishable patterns (data not shown¹³). The two homologous, hexameric proteins also show comparable progressions of digestion. In both enzymes, the first polypeptide produced is generated from a loss of the C terminus. Interestingly, a trypsin digestion of the T4 gp41, the replicative DNA helicase in the T4 system, has an initial cleavage in this same region.³⁴

The N-terminal region of T7 gp4 comprises two structural domains, a ZBD and a RPD that are separated by a non-structured tether.^{16,17} This tether region is the next site of cleavage and results in a loss of the ZBD. Digestion of the human mtDNA helicase follows this pattern, with proteolytic cleavage at an accessible homologous site in its N-terminal region. In contrast, we were unable to demonstrate an N-terminal cleavage product corresponding to the ZBD of T7 gp4 despite several attempts employing different strategies, and thus find no evidence that it exists as a stable product. However, we found that the trypsin-sensitive site at aa R275 in T7 gp4 reported by Washington *et al.* corresponds to a cleavage occurring at the highly conserved R371 in the human mtDNA helicase.¹⁵ This is consistent with the presence of a region that is structurally homologous to the linker region of the bacteriophage protein.

Our sequence analysis of the proteolytic products of the human mtDNA helicase also reveals similarities to another member of the hexameric family of

replicative helicases, *E. coli* DnaB protein. Under limited conditions, a trypsin digestion of DnaB protein produces N and C-terminal structural domains from a single cleavage occurring in a proteolytically sensitive region.¹⁴ The trypsin digestions in our study also generated N and C-terminal polypeptides whose cleavage sites correspond to predicted unstructured regions of the intact protein. Based on the digestion pattern observed with DnaB protein, we would propose that the T30 and T28 polypeptides of human mtDNA helicase result from cleavage of T57 at a position C-terminal to aa 371. However, it is not clear how this relates to the alternate production of T50 and T34 from T57. Without knowing the precise C terminus of T28, we are unable to confirm a hypothesis of alternative proteolytic pathways, and cannot at this time demonstrate the pathway leading to fragment T28. We speculate that multiple proteolytic patterns are indicative of either conformational changes or asymmetry within the hexamer. Regardless, experiments using multiple proteases and product stability analyses using a time-course of digestion document that the minimal polypeptides T30 and T28 represent distinct structural domains, and establish the modular architecture of the protein.

The majority of the T7 gp4 oligomerization properties are due to interactions between helix A and helices D1–D3.^{21,22,26} We show that a hexameric form is still retained in T50 but lost in T28, which exists as a monomer. This suggests that the region containing the residues responsible for maintaining a stable oligomer lies within aa 405–590, and that this region encompasses a homologous region comprising helices D1–D3 in the bacteriophage protein. Although T28 contains the predicted helix A, it lacks helices D1–D3 and therefore would be incapable of the necessary interactions yielding a stable hexamer. This is consistent with findings reported for the monomeric primase region of bacteriophage T7 gp4.¹⁸

Studies of the bacteriophage T7 and T4 replicative DNA helicases have revealed the importance of their C termini for protein–protein interactions within the replisome.^{34,35} In particular, this region in T7 gp4 provides a regulatory role involved in coordinating leading and lagging DNA strand synthesis. However, in neither bacteriophage system has the C terminus been shown to affect nucleotide hydrolysis. Interestingly, we show a significant increase in nucleotide hydrolysis for the recombinant P66 protein lacking its C terminus. Thus, the C terminus of the full-length enzyme apparently provides a negative regulatory function. In contrast, nucleotide hydrolysis is decreased to full-length enzyme activity in the recombinant P57 protein that also lacks its N terminus. Our data suggest that the N terminus may contribute to nucleotide hydrolysis in the full-length protein, similar to *Methanobacterium thermoautotrophicum* MCM (mtMCM), an enzyme that serves as a model for eukaryotic systems.³⁶ Alternatively, the loss in activity due to removal of the N terminus may result from a deleterious

conformational change. In contrast however, an *E. coli* DnaB fragment lacking its N terminus retains the nucleotide hydrolysis activity of the full-length enzyme.¹⁴ In T7 gp4, the N-terminal region contains the ZBD, which contacts DNA and pauses the enzyme during primer synthesis.³⁷ Furthermore, a crystal structure of *mtMCM* shows that this region in the archaeal replicative helicase is involved in dodecamerization *via* zinc motifs.³⁸ The contribution of the C terminus to the activity of the mtDNA helicase is novel among the replicative helicases, and the potential involvement of the N terminus in this function suggests the possibility of cellular functions extending beyond its role in the elongation phase of DNA replication.

We found that a hexameric P57 exhibits higher ATPase activity as compared to heterohexameric T57/T50, and note that T50 lacks residues critical to the nucleotide hydrolysis reaction. In particular, we propose that the reduced activity results from the loss of a putative arginine finger: R609 of the human mtDNA helicase likely serves a role analogous to R522 in T7 gp4²⁰ (Figure 3(b)). The arginine finger, found in all ring helicases, is responsible for ligating the γ -phosphate during the nucleotide hydrolysis reaction. Upon hydrolysis, the release of the γ -phosphate displaces the arginine finger, leading to a conformational change and subsequent DNA translocation.^{11,20,26} The crystal structure of the T7 gp4 demonstrates that within the hexamer, subunits A and B are rotated 15° with respect to each other, and the C subunits are rotated 30° in the opposite direction to compensate for the torsional strain produced.²⁰ This asymmetric orientation puts the arginine finger in the optimal position for binding the γ -phosphate in the neighboring A and B subunits, whereas in the C subunits it is too far away.²⁰ In the DnaB-like helicases, residues at the subunit interface of one monomer bind the nucleotide, but it is the arginine finger of the neighboring monomer that participates in ligating the γ -phosphate.^{20,26} Our proteolysis data suggest that the human mtDNA helicase may also have an asymmetric orientation within the hexamer. In the T57/T50 heterohexamer, catalytically competent hexamers would likely comprise oligomers in which T57 occupies minimally the four neighboring subunits that are oriented to bind nucleotide. Although T50 is less abundant in our preparations, hexamers in which it neighbors T57 would be functionally inactive.

We found that the N-terminal proteolytic product T28 lacks ATPase activity. T28 does not contain the classic Walker A and Walker B motifs located in conserved motifs I and II in SF-1 and SF-2 family members, like the N-terminal only domains of T7 gp4 and *E. coli* DnaB protein that exist as monomers under certain conditions.^{4,5,14,16,24} Clearly, the ATPase activity of the mitochondrial enzyme requires residues within the C-terminal domain. At the same time, similarities shared by T28 and the N-terminal fragments of T7 gp4 and DnaB protein may provide insight for its possible role in protein-protein interactions and/or helicase regulation.²⁵

Bacteriophage T7 gp4 and *E. coli* DnaB protein are members of the DnaB-like family of replicative DNA helicases.^{4,9,33} Here we probe the physical features of the mitochondrial enzyme, and conclude that the hexameric human mtDNA helicase is a modular protein comprising two N and C-terminal structural domains. We also show that the essential residues for nucleotide hydrolysis, as well as those involved in oligomerization, reside exclusively in the C-terminal domain. Together with previous functional data, the human mtDNA helicase may be classified firmly within the DnaB-like family.³

Materials and Methods

Materials

Proteases, soybean trypsin inhibitor (STI), bovine serum albumin, bovine liver catalase, rabbit muscle L-lactate dehydrogenase, and yeast alcohol dehydrogenase were purchased from Sigma. Color markers for SDS-PAGE were also purchased from Sigma. [α -³²P]ATP was purchased from MP Biosciences. *E. coli* DnaB protein was a gift from Dr Jon Kaguni of this department. His-tagged monoclonal antibody was purchased from Becton Dickinson Biosciences. Polyethyleneimine chromatography paper impregnated with fluorescent indicator was purchased from Brinkman Instruments, Inc. Amphotericin, penicillin-G, streptomycin and phenylmethylsulfonyl fluoride (PMSF) were purchased from Sigma. PMSF was prepared as a 0.1 M stock solution in isopropyl alcohol and stored at -20 °C. Sodium metabisulfite (J. T. Baker Chemical Co.) was prepared as a 1 M stock solution at pH 7.5 and stored at -20 °C. 1,4-dithio-DL-threitol was purchased from Research Organics. Leupeptin was purchased from the Peptide Institute, Minoh-Shi, Japan, and was prepared as a 1 mg/ml stock solution in 50 mM Tris-HCl (pH 8.0), 2 mM EDTA and stored at -20 °C. SDS for gel electrophoresis was from Pierce. Nickel-nitrilotriacetic acid resin was purchased from Qiagen. *Spodoptera frugiperda* cells were a gift from Dr Suzanne Thiem. TC-100 insect cell culture medium and fetal bovine serum were from Life Technologies, Inc. Insect cell transfection buffer and Grace's medium were from PharMingen. Baculoviruses encoding N and C-terminally His-tagged human mtDNA helicase lacking its mitochondrial targeting sequence (aa 43–684) were a gift from the Falkenberg laboratory at the Karolinska Institute. Baculoviral plasmid DNA encoding N-terminally His-tagged human mtDNA helicase P66 (aa 43–633) and N-terminally His-tagged human mtDNA helicase P57 (aa 146–633) were constructed by and purchased from ATG Laboratories, Inc. A Superdex 200 HR 10/30 column was purchased from Amersham Biosciences.

Construction of recombinant baculoviruses P66 and P57

Linearized wild-type baculovirus DNA (BaculoGold, 0.5 μ g; PharMingen) and purified transfer plasmid DNAs (ATG Laboratories) encoding P66 or P57 human mtDNA helicase were co-transfected in transfection buffer containing 25 mM Hepes (pH 7.1), 125 mM CaCl₂, 140 mM NaCl for 4 h at 27 °C following the manufacturer's recommendations. Recombinant viruses were amplified in *S.*

frugiperda cells to titers of approximately 2×10^8 plaque forming units/ml.

Protein overexpression and purification of human mtDNA helicase

S. frugiperda cells were grown in TC-100 insect culture cell medium containing 10% (v/v) fetal bovine serum and infected with recombinant viruses at a multiplicity of infection of 5 at 27 °C. Cells were harvested at 72 h post-infection by centrifugation and washed with an equal volume of cold phosphate-buffered saline, recentrifuged, frozen in liquid nitrogen and stored at -80 °C. The frozen cell pellets were thawed on ice, and all further steps were performed at 0–4 °C. Cells were suspended in 1/45 volume of original cell culture in 25 mM Tris-HCl (pH 8.0), 1 mM PMSF, 10 mM sodium metabisulfite, 2 µg/ml leupeptin, 10 mM 2-mercaptoethanol. Cells were lysed in a Dounce homogenizer using 20 strokes of the tight pestle. The homogenate was adjusted to a final salt concentration of 1 M NaCl followed by centrifugation. The soluble extract (Fr I) was diluted with an equal volume of 50 mM Tris-HCl (pH 8.0), 0.6 M NaCl, 10% (v/v) glycerol and loaded onto a nickel-nitrilotriacetic acid agarose column (2.5 ml resin/liter of cells) equilibrated with buffer containing 50 mM Tris-HCl (pH 8.0), 0.6 M NaCl, 10% glycerol, 10 mM imidazole, 1 mM PMSF, 10 mM sodium metabisulfite, 2 µg/ml leupeptin, 10 mM 2-mercaptoethanol. The column was washed with the equilibration buffer containing 0.6 M NaCl and 10 mM imidazole, and the bound protein was eluted with buffers containing 25 mM, 250 mM and 500 mM imidazole. Recombinant proteins eluted at 250 mM imidazole; fractions were analyzed by SDS-PAGE and silver staining, and fractions were pooled accordingly as Fr II. Fr II was diluted to an ionic equivalent of 150 mM NaCl and loaded onto a heparin Sepharose column (2.7 mg/ml resin) equilibrated with buffer containing 20 mM Tris-HCl (pH 7.5), 150 mM NaCl, 10% glycerol, 0.5 mM EDTA, 1 mM PMSF, 10 mM sodium metabisulfite, 2 µg/ml leupeptin, 10 mM 2-mercaptoethanol. The column was washed in the equilibration buffer containing 200 mM NaCl and the bound protein was eluted with buffers containing 0.6 and 1 M NaCl. Recombinant proteins eluted at 350–700 mM NaCl; fractions were analyzed by SDS-PAGE followed by silver staining, and pooled accordingly as Fr III. Fr III was loaded onto 12–30% glycerol gradients in 35 mM Tris-HCl (pH 7.5), 330 mM NaCl, 1 mM EDTA, 1 mM PMSF, 10 mM sodium metabisulfite, 2 µg/ml leupeptin, 10 mM 2-mercaptoethanol. Gradients were centrifuged at 37,000 rpm for 30 h at 4 °C in a Beckman SW41 rotor. Fractions were analyzed by SDS-PAGE and silver staining, and pooled accordingly as Fr IV. Fr IV was frozen in liquid nitrogen and stored in aliquots at -80 °C.

Analytical glycerol gradient sedimentation of human mtDNA helicase

N and C-His-tagged human mtDNA helicase (Fr IV; 15 µg, 25 µg/ml) were layered onto preformed 12–30% glycerol gradients (10 ml) containing 35 mM Tris-HCl (pH 7.5), 330 mM NaCl, 1 mM EDTA, 1 mM PMSF, 10 mM sodium metabisulfite, 2 µg/ml leupeptin, 10 mM 2-mercaptoethanol. Protein standards of known *S* values run in parallel were: jack bean urease, 18.6 *S*; bovine liver catalase, 11.3 *S*; rabbit muscle L-lactate dehydrogenase, 7.3 *S*. Centrifugation was at 37,000 rpm for 30 h at 4 °C in a Beckman SW41 rotor. Fractions were analyzed by SDS-

PAGE and silver staining. The data were plotted as *S* value versus fraction number to obtain a sedimentation coefficient for the human mtDNA helicase.

Gel filtration

C-terminal His-tagged human mtDNA helicase (200 µg, 0.4 µg/ml) was chromatographed on a Superdex 200 HR 10/30 column (10 mm × 300–310 mm, Amersham Biosciences) equilibrated with buffer containing 35 mM Tris-HCl (pH 7.5), 330 mM NaCl, 8% glycerol, 1 mM EDTA at a flow rate of 0.25 ml/min at 4 °C. The column was calibrated with chymotrypsinogen A ($R_s = 28$ Å), hen egg ovalbumin ($R_s = 30$ Å), yeast alcohol dehydrogenase ($R_s = 46$ Å), bovine liver catalase ($R_s = 52$ Å) and bovine thyroid thyroglobulin ($R_s = 85$ Å), where R_s is the Stoke's radius. Fractions (0.25 ml) were collected and aliquots were examined by SDS-PAGE and silver staining to confirm ultraviolet trace recordings.

Limited proteolysis

N and C-terminally His-tagged human mtDNA helicase (2 µg) was proteolyzed in 50 mM Tris-HCl (pH 8.0), 50% glycerol, 4 mM MgCl₂, 2 mM 2-mercaptoethanol for 10 min at 20 °C. The trypsin titration contained a trypsin concentration range of 0.05 µg–10.0 µg. The multiple protease digestion used optimal helicase:protease ratios (w/w) as follows: trypsin 1:1, chymotrypsin 1:0.35, thermolysin 1:1 (for 40 min), subtilisin 1:0.005, papain 1:0.18. The trypsin time-course digestion used a trypsin: helicase ratio of 1:1. At various time points indicated in the legend to Figure 4, 30 µl aliquots were removed from the digested sample. All trypsin digestions were stopped by the addition of STI at a STI:trypsin ratio of 2:1. All of the other protease digestions were stopped by addition of 10 mM sodium metabisulfite, 2 µg/ml leupeptin, and 1% SDS. Digestion products were separated using SDS-PAGE, and analyzed by silver staining and/or immunoblotting.

N-terminal sequencing

Following trypsin digestions designed to optimize production of individual fragments, proteolytic products were separated by SDS-PAGE and transferred onto a polyvinylidene difluoride membrane in CAPS buffer at a constant current of 150 mAmp for 16 h. The membranes were stained with 0.05% R250 Coomassie Blue followed by destaining in 40% methanol. The proteolytic fragments were visualized, cut from the membrane and submitted for N-terminal sequencing analysis by standard Edman chemistry.

Preparation of proteolytic products for gel filtration analysis

To produce T57/T50, C-terminally His-tagged mtDNA helicase was proteolyzed at a helicase:trypsin ratio of 1:2 for 10 min at 20 °C, and the digestion was stopped with STI. To produce T34/T28, C-terminally His-tagged mtDNA helicase was proteolyzed at a helicase:trypsin ratio of 1:0.25 for 55 min at 20 °C and stopped with STI. T57/T50 and T34/T28 were chromatographed on the Superdex 200 HR 10/30 column and analyzed as for the full-length protein.

ATPase assay

Reaction mixtures (20 μ l) contained 20 mM Tris-HCl (pH 7.5), 4 mM MgCl₂, 0.1 mg/ml bovine serum albumin, 10% glycerol, 0.5 mM ATP, 10 mM dithiothreitol, 4 μ Ci of [α -³²P]ATP, 100 μ M DNase I-activated calf thymus DNA, and N and C-terminally His-tagged full-length mtDNA helicase, T66/T57, T57/T50, P66, P57 and T28. Incubation was for 15 min at 37 °C unless indicated otherwise. The reactions were stopped by addition of EDTA to 20 mM and by placing the reaction tubes on ice, and aliquots (2 μ l) were spotted onto Polygram polyethyleneimine cellulose paper with 0.5 μ l of 50 mM ADP/ATP. The Polygram polyethyleneimine cellulose paper was developed in 1 M formic acid, 0.5 M LiCl. The position of both ADP and ATP bands was visualized by ultraviolet light. The corresponding bands were isolated and the radioactivity was measured by liquid scintillation counting.

Acknowledgements

We thank members of Dr Jon Kaguni's laboratory for critical discussions of the data. This work was supported by National Institutes of Health Grant GM45295 to L.S.K.

References

- Tyynismaa, H., Sembongi, H., Bokori-Brown, M., Granycome, C., Ashley, N., Poulton, J. *et al.* (2004). Twinkle helicase is essential for mtDNA maintenance and regulates mtDNA copy number. *Hum. Mol. Genet.* **13**, 3219–3227.
- Spelbrink, J. N., Li, F. Y., Tiranti, V., Nikali, K., Yuan, Q. P., Tariq, M. *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. *Nature Genet.* **28**, 223–231.
- Korhonen, J. A., Gaspari, M. & Falkenberg, M. (2003). TWINKLE has 5' \rightarrow 3' DNA helicase activity and is specifically stimulated by mitochondrial single-stranded DNA-binding protein. *J. Biol. Chem.* **278**, 48627–48632.
- Caruthers, J. M. & McKay, D. B. (2002). Helicase structure and mechanism. *Curr. Opin. Struct. Biol.* **12**, 123–133.
- Walker, J. E., Saraste, M., Runswick, M. J. & Gay, N. J. (1982). Distantly related sequences in the α - and β -subunits of ATP synthase, myosin, kinases and other ATP-requiring enzymes and a common nucleotide binding fold. *EMBO J.* **1**, 945–951.
- Egelman, E. H., Yu, X., Wild, R., Hingorani, M. M. & Patel, S. S. (1995). Bacteriophage T7 helicase/primase proteins form rings around single-stranded DNA that suggest a general structure for hexameric helicases. *Proc. Natl Acad. Sci. USA*, **92**, 3869–3873.
- Bujalowski, W., Klonowska, M. M. & Jezewska, M. J. (1994). Oligomeric structure of *Escherichia coli* primary replicative helicase DnaB protein. *J. Biol. Chem.* **269**, 31350–31358.
- Dong, F., Gogol, E. P. & von Hippel, P. H. (1995). The phage T4-coded DNA replication helicase (gp41) forms a hexamer upon activation by nucleoside triphosphate. *J. Biol. Chem.* **270**, 7462–7473.
- LeBowitz, J. H. & McMacken, R. (1986). The *Escherichia coli* dnaB replication protein is a DNA helicase. *J. Biol. Chem.* **261**, 4738–4748.
- Scherzinger, E., Ziegelin, G., Barcena, M., Carazo, J. M., Lurz, R. & Lanka, E. (1997). The RepA protein of plasmid RSF1010 is a replicative DNA helicase. *J. Biol. Chem.* **272**, 30228–30236.
- Niedenzu, T., Roleke, D., Bains, G., Scherzinger, E. & Saenger, W. (2001). Crystal structure of the hexameric replicative helicase RepA of plasmid RSF1010. *J. Mol. Biol.* **306**, 479–487.
- Yu, X., Jezewska, M. J., Bujalowski, W. & Egelman, E. H. (1996). The hexameric *E. coli* DnaB helicase can exist in different quaternary states. *J. Mol. Biol.* **259**, 7–14.
- Bird, L. E., Hakansson, K., Pan, H. & Wigley, D. B. (1997). Characterization and crystallization of the helicase domain of bacteriophage T7 gene 4 protein. *Nucl. Acids Res.* **25**, 2620–2626.
- Nakayama, N., Arai, N., Kaziro, Y. & Arai, K. (1984). Structural and functional studies of the dnaB protein using limited proteolysis. Characterization of domains for DNA-dependent ATP hydrolysis and for protein association in the primosome. *J. Biol. Chem.* **259**, 88–96.
- Washington, M. T., Rosenberg, A. H., Griffin, K., Studier, F. W. & Patel, S. S. (1996). Biochemical analysis of mutant T7 primase/helicase proteins defective in DNA binding, nucleotide hydrolysis, and the coupling of hydrolysis with DNA unwinding. *J. Biol. Chem.* **271**, 26825–26834.
- Frick, D. N., Baradaran, K. & Richardson, C. C. (1998). An N-terminal fragment of the gene 4 helicase/primase of bacteriophage T7 retains primase activity in the absence of helicase activity. *Proc. Natl Acad. Sci. USA*, **95**, 7957–7962.
- Kato, M., Ito, T., Wagner, G., Richardson, C. C. & Ellenberger, T. (2003). Modular architecture of the bacteriophage T7 primase couples RNA primer synthesis to DNA synthesis. *Mol. Cell*, **11**, 1349–1360.
- Kato, M., Frick, D. N., Lee, J., Tabor, S., Richardson, C. C. & Ellenberger, T. (2001). A complex of the bacteriophage T7 primase-helicase and DNA polymerase directs primer utilization. *J. Biol. Chem.* **276**, 21809–21820.
- Kato, M., Ito, T., Wagner, G. & Ellenberger, T. (2004). A molecular handoff between bacteriophage T7 DNA primase and T7 DNA polymerase initiates DNA synthesis. *J. Biol. Chem.* **279**, 30554–30562.
- Singleton, M. R., Sawaya, M. R., Ellenberger, T. & Wigley, D. B. (2000). Crystal structure of T7 gene 4 ring helicase indicates a mechanism for sequential hydrolysis of nucleotides. *Cell*, **101**, 589–600.
- Guo, S., Tabor, S. & 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.
- Lee, S. J. & 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.
- Lu, Y. B., Ratnakar, P. V., Mohanty, B. K. & Bastia, D. (1996). Direct physical interaction between DnaG primase and DnaB helicase of *Escherichia coli* is necessary for optimal synthesis of primer RNA. *Proc. Natl Acad. Sci. USA*, **93**, 12902–12907.
- Biswas, S. B., Chen, P. H. & Biswas, E. E. (1994).

- Structure and function of *Escherichia coli* DnaB protein: role of the N-terminal domain in helicase activity. *Biochemistry*, **33**, 11307–11314.
25. Fass, D., Bogden, C. E. & Berger, J. M. (1999). Crystal structure of the N-terminal domain of the DnaB hexameric helicase. *Structure*, **7**, 691–698.
26. Sawaya, M. R., Guo, S., Tabor, S., Richardson, C. C. & Ellenberger, T. (1999). Crystal structure of the helicase domain from the replicative helicase-primase of bacteriophage T7. *Cell*, **99**, 167–177.
27. Toth, E. A., Li, Y., Sawaya, M. R., Cheng, Y. & Ellenberger, T. (2003). The crystal structure of the bifunctional primase-helicase of bacteriophage T7. *Mol. Cell*, **12**, 1113–1123.
28. Story, R. M. & Steitz, T. A. (1992). Structure of the recA protein-ADP complex. *Nature*, **355**, 374–376.
29. Velankar, S. S., Soutanas, P., Dillingham, M. S., Subramanya, H. S. & Wigley, D. B. (1999). Crystal structures of complexes of PcrA DNA helicase with a DNA substrate indicate an inchworm mechanism. *Cell*, **97**, 75–84.
30. Lenzen, C. U., Steinmann, D., Whiteheart, S. W. & Weis, W. I. (1998). Crystal structure of the hexamerization domain of N-ethylmaleimide-sensitive fusion protein. *Cell*, **94**, 525–536.
31. Richardson, J. P. (1996). Structural organization of transcription termination factor Rho. *J. Biol. Chem.* **271**, 1251–1254.
32. Yu, R. C., Hanson, P. I., Jahn, R. & Brunger, A. T. (1998). Structure of the ATP-dependent oligomerization domain of N-ethylmaleimide sensitive factor complexed with ATP. *Nature Struct. Biol.* **5**, 803–811.
33. Patel, S. S. & Hingorani, M. M. (1993). Oligomeric structure of bacteriophage T7 DNA primase/helicase proteins. *J. Biol. Chem.* **268**, 10668–10675.
34. Richardson, R. W. & Nossal, N. G. (1989). Trypsin cleavage in the COOH terminus of the bacteriophage T4 gene 41 DNA helicase alters the primase-helicase activities of the T4 replication complex *in vitro*. *J. Biol. Chem.* **264**, 4732–4739.
35. Notarnicola, S. M., Mulcahy, H. L., Lee, J. & Richardson, C. C. (1997). The acidic carboxyl terminus of the bacteriophage T7 gene 4 helicase/primase interacts with T7 DNA polymerase. *J. Biol. Chem.* **272**, 18425–18433.
36. Chong, J. P., Hayashi, M. K., Simon, M. N., Xu, R. M. & Stillman, B. (2000). A double-hexamer archaeal minichromosome maintenance protein is an ATP-dependent DNA helicase. *Proc. Natl Acad. Sci. USA*, **97**, 1530–1535.
37. Lee, J. B., Hite, R. K., Hamdan, S. M., Xie, X. S., Richardson, C. C. & van Oijen, A. M. (2006). DNA primase acts as a molecular brake in DNA replication. *Nature*, **439**, 621–624.
38. Fletcher, R. J., Bishop, B. E., Leon, R. P., Sclafani, R. A., Ogata, C. M. & Chen, X. S. (2003). The structure and function of MCM from archaeal *M. Thermoautotrophicum*. *Nature Struct. Biol.* **10**, 160–167.

Edited by J. Doudna

(Received 14 December 2006; accepted 19 January 2007)
Available online 7 February 2007