Quaternary Polymorphism of Replicative Helicase G40P: Structural Mapping and Domain Rearrangement

Rafael Núñez-Ramírez¹, Yolanda Robledo¹, Pablo Mesa², Silvia Ayora²
Juan Carlos Alonso², José María Carazo¹* and Luis Enrique Donate¹

1 Department of Macromolecular Structure, Centro Nacional de Biotecnología CSIC, Campus Universidad Autónoma de Madrid, 28049 Cantoblanco, Madrid, Spain
2 Department of Microbial Biotechnology, Centro Nacional de Biotecnología, CSIC, Campus Universidad Autónoma de Madrid, 28049 Cantoblanco Madrid, Spain

Quaternary polymorphism is a distinctive structural feature of the DnaB family of replicative DNA hexameric helicases. The Bacillus subtilis bacteriophage SPP1 gene 40 product (G40P) belongs to this family. Three different quaternary states have been described for G40P homohexamers, two of them with C₃ symmetry, and the other with C₆ symmetry. We present three-dimensional reconstructions of the different architectures of G40P hexamers and a variant lacking the N-terminal domain. Comparison of the G40P and the deletion mutant structures sheds new light on the functional roles of the N and C-terminal domains, at the same time that it allows the direct structural mapping of these domains. Based on this new information, hybrid EM/X-ray models are presented for all the different symmetries. These results suggest that quaternary polymorphism of hexameric helicases may be implicated in the translocation along the DNA.

0022-2836/$ - see front matter © 2006 Elsevier Ltd. All rights reserved.

Introduction

DNA helicases are molecular motors that unwind double-stranded DNA fuelled by the energy produced by nucleoside triphosphate hydrolysis.¹ Helicases are involved in all processes of DNA metabolism that need single-stranded DNA as a substrate. Sequence analysis of this group of enzymes has led to their classification in different superfamilies (SF1 to SF3) and several smaller families (e.g. DnaB family), which have in common a set of conserved helicase motifs.²

Helicases assemble into different oligomeric states. Structural studies of monomeric,³,⁴ dimeric⁵ and hexameric helicases⁶,⁷ have shown that conserved sequence motifs of these enzymes are situated in the nucleotide-binding core which was first described for the RecA protein.⁸ Moreover, the most conserved residues within the characteristic motifs of helicases are implicated in nucleotide binding and NTPase activity, all of them around the so-called P-loop. A conserved arginine finger (Arg-finger) is also essential for NTPase activity.⁹,¹⁰ This Arg-finger belongs to another subunit in the case of oligomeric helicases or to another domain in monomeric helicases.¹¹

The helicases of the DnaB family are intimately associated with the DNA primase and participate in the replicative process assembled as hexamers. The main representative of this family is Escherichia coli DnaB (EcoDnaB). Other members are gp4 of E. coli T7 bacteriophage (T7gp4) and G40P of Bacillus subtilis SPP1 bacteriophage.¹² Both electron microscopy (EM) and X-ray diffraction studies¹² have shown that hexameric helicases assemble as ring-shaped structures with a central channel. Interaction with DNA seems to be through their inner channel,¹³–¹⁵ and one of the conserved motifs (motif H4) has been described to be responsible for this interaction.¹⁶ Some helicases of the DnaB family have previously been shown to assemble in several states with different point symmetry. Projection images of EcoDnaB¹⁷ and G40P¹⁸ revealed the existence of particles with 3-fold (C₃) and 6-fold rotational symmetry (C₆), as well as an intermediate state. Replicative helicases of other families, for example RepA of plasmid RSF1010¹⁹ and papilloma E1 helicase,²⁰ may display a similar polymorphism, although the intermediate state was not identified.

The RecA-like helicase domains are usually joined to other domains. For example, limited trypsinolysis studies have demonstrated that
Eco-DnaB protein (52 kDa) could be cleaved into two major discrete polypeptides, to render a 33 kDa C-terminal domain (CTD) and a 12 kDa N-terminal domain (NTD), linked by a small hinge region.21,22 Whereas the helicase domain laid in the CTD, the NTD has an undetermined associated function. In a similar way, the bacteriophage T7 gp4 contains a helicase domain physically linked to a primase domain.23 This modular organization may be common in helicases.

The G40P helicase is required for the initiation of B. subtilis bacteriophage SPP1 DNA replication.24 G40P hexamers show all the activities associated with DnaB-like enzymes, including DNA unwinding with a 5' to 3' polarity.25 Analysis of the G40P primary structure indicates that this protein presents a modular organization similar to EcoDnaB, with a NTD and a CTD linked by a hinge region. A mutant lacking the NTD (G40PΔN109) retains the DNA binding and the DNA helicase activity, although without the characteristic functional polarity of the wild-type (wt) enzyme. In a pH-dependent manner, G40PΔN109 displays helicase activity in both polarities, suggesting that the NTD main role is the regulation of the enzyme helicase action.

We have generated three-dimensional (3D) reconstructions for the different states of G40P helicase, as well as for the G40PΔN109 mutant. Our study shows that G40PΔN109 primarily oligomerizes as a hexamer with preferential 6-fold rotational symmetry, although it is able to undergo the structural reorganization required for the transition between the different symmetry states. We present the first 3D data for a hexameric helicase of Gram-positive bacterial origin, allowing the experimental localization of the NTD within the whole structure of a DnaB-like replicative helicase. The NTD plays an active role in the stability of the complex and in the symmetry transition, but is not essential for enzyme activity. A model for the rearrangement of the CTD and NTD in the symmetry transition, as well as their functional role, is proposed.

Results

Three-dimensional structure of G40P

3D reconstructions were obtained for the three different architectures of G40P. The Random Conical Tilt approach was followed to obtain non-biased 3D initial references. Figure 1(a) and (b) shows a representative tilting pair for a G40P preparation in buffer A. G40P oligomers show a preferred view in which the protein appears as a globular-shaped particle with a central stain-penetrating region. Non-overlapping pairs of particles were manually selected and extracted from the micrographs to a total of 13,743. The set of untitled images were used for the initial 2D alignment and classification, whereas the corres-

![Figure 1. EM and 2D analysis of G40P hexamers. (a) Untilted and (b) 50° tilted electron micrographs of negatively stained G40P oligomers in buffer A. Arrowheads point to some representative particles; numbers indicate correspondence between partners in the tilting pair. (c), (e) and (g) Rotational power spectra, calculated between radii 5–8 nm, of C₃, C₃C₆, and C₆ average images, respectively. (d), (f) and (h) 2D average images of the re-aligned C₃, C₃C₆ and C₆ groups, respectively. The bar represents 50 nm in (a) and 5 nm in (d). Arrowheads in (f) indicate the main differences between the C₃ and C₃C₆ average images.](image)
The unaligned images were aligned and classified as described in Materials and Methods. Classification was performed using the rotational power spectra rather than the images themselves. Classes with rotational symmetry similar to those identified in previous studies were selected. Rotational spectra representative of each of these groups are presented in Figure 1(c), (e) and (g). The C3 group is characterized by a rotational power spectrum that shows a maximum value for the 3-fold component followed by a noticeable contribution of the 6-fold one (Figure 1(c)). The rotational power spectrum of the C6 group only shows a peak for the 6-fold harmonic (Figure 1(g)). The rotational spectrum of the third group shows the 6-fold harmonic as the strongest one, and an important contribution from the 3-fold component (Figure 1(e)), as was described for the intermediate state. This group will be referred to as the C3C6 symmetry group.

Only particles clearly corresponding to the described symmetry patterns were selected. The three groups were roughly populated with 700 particles each. Images associated to each group were independently realigned. The average images calculated for each group are shown in Figure 1(d), (f) and (h) next to their corresponding rotational power spectra. As previously reported, the average images from groups C3 (Figure 1(d)) and C3C6 (Figure 1(f)) present six central masses and three smaller blobs in the vertexes, while the average image from group C6 (Figure 1(h)) assembles as six masses with the same shape and orientation. The main difference observed between the average images of the C3 and C3C6 classes concerns the presence of a weak density protruding out of the three central masses (arrowheads in Figure 1(f)). This restrictive selection is due to the fact that the Random Conical Tilt approach strictly requires that all the unaligned images, partners of the tilted images used for the 3D reconstruction, were structurally homogeneous and in exactly the same orientation with respect to the supporting film. Therefore, only images that clearly belonged to one of the symmetry groups and did not present any sign of possible rocking were further considered for 3D reconstruction.

The angular information derived from the rotational alignment of the unaligned images, together with the data from the tilting process, were used to define the conical geometry in order to obtain an initial 3D reconstruction for each group. Each initial reconstruction was then used as the starting point in an iterative process of angular refinement performed only within their associated data set of images, until convergence was reached.

Figure 2 presents the isosurface rendering of the three G40P reconstructed volumes. The three reconstructions are of similar height (5.5 nm) and possess a passing channel that runs open from top to bottom with an even diameter of roughly 4 nm. The C3 and C3C6 reconstructions share a very similar architecture. The maps present three main units that can be subdivided into two masses. In other words, they are consistent with a trimer of dimers. Rotational analysis of sections taken perpendicularly to the rotational axis of the volumes (data not shown) revealed that the C3 and C3C6 volumes possess two different faces, one of predominant 3-fold symmetry (bottom face) and an opposite face with 6-fold symmetry (top face). Similar results were observed with the homologous EcoDnaB helicase. Despite the overall similarity between the C3 and C3C6 volumes, it is worth noting that there are some minor but very reproducible differences. The main one lies in a small protrusion sticking out in three of the six monomers (arrowhead in Figure 2), that probably corresponds to the small protrusion mentioned in the 2D analysis (arrowhead in Figure 1(f)). Conversely, the C6 reconstruction (Figure 2, column C6) shows a hexamer whose architecture is based on six identical masses rendering a volume with 6-fold symmetry on both faces. Nevertheless, the top and bottom faces are easily distinguished, since the top face shows a wider diameter with the six monomers, of isolated appearance, reaching out towards the exterior. The diameter of the bottom face is narrower and the neighboring monomers make closer contacts.

Three-dimensional structure of G40PΔN109

Previously it was shown that the G40PΔN109 enzyme, which contains the hinge region and the CTD, but lacks the NTD, is able to unwind DNA. We used here G40PΔN109 to localize the NTD in wt G40P structure by 3D-EM studies. Figure 3(a) and (b) presents a representative tilting pair of negatively stained G40PΔN109 in buffer A. Globular-shaped particles of G40PΔN109 show morphology and dimensions similar to the wild-type particles. The quality of the G40PΔN109 images (in terms of contrast and sharpness) was systematically lower than those of G40P. Also, the number of particles was always lower in the G40PΔN109 samples, although both samples were diluted to the same concentration. A total of 4195 non-overlapping pairs of particles were manually selected and processed as described for G40P.

Unlike G40P, classification showed no groups of predominant 3-fold symmetry. A group of 1500 particles was selected, whose spectra had a single peak of energy at the 6-fold component (Figure 3(c)). The average image for this C6 group (Figure 3(d)) shows a very similar appearance to the average image of the C6 group of G40P (Figure 1(h)), indicating that G40PΔN109 oligomerizes as a hexamer. This result also suggests that the CTD and NTD are on top of each other along the rotational axis, because there is no clear reduction in the dimension of the particle.
The Random Conical Tilt approach was used to calculate an initial volume, using only the set of tilted images assessed to be C₆. This volume was refined through an iterative projection matching algorithm. A final volume was then obtained using all the images collected for this mutant, because G40PΔN109 oligomers do not seem to exhibit quaternary polymorphism under the buffer A conditions. This procedure ensures a more complete coverage of the projection space than the one typically obtained by using only the tilted images.

Figure 2. 3D structures of the three G40P architectures. Isosurface renderings at threshold values accounting for 100% of the mass are shown for the C₃ (left column), C₃C₆ (middle column) and C₆ (right column) groups. Reconstructions are shown from different viewing points (top, capsized and side on top, middle and bottom rows, respectively). Top and capsized views are related by a 45° rotation, and side and top views by a 90° rotation. Arrowheads mark the main differences between the very similar C₃ and C₃C₆ structures. The maximum diameter is 14 nm and the height is roughly 5.5 nm for all structures. The estimated resolution is 24 Å in all instances.

The Random Conical Tilt approach was used to calculate an initial volume, using only the set of tilted images assessed to be C₆. This volume was refined through an iterative projection matching algorithm. A final volume was then obtained using all the images collected for this mutant, because G40PΔN109 oligomers do not seem to exhibit quaternary polymorphism under the buffer A conditions. This procedure ensures a more complete coverage of the projection space than the one typically obtained by using only the tilted images.

Figure 3. EM and 2D analysis of G40PΔN109 hexamers. (a) Untilted and (b) 50° tilted electron micrographs of negatively stained G40PΔN109 oligomers in buffer A. Arrowheads point to some representative particles; numbers indicate correspondence between partners in the tilting pair. (c) Rotational power spectrum, calculated between radii 5–8 nm of the particle, representative of images classified as C₆, the only group found. (d) 2D average of the re-aligned C₆ group. The bar represents 50 nm in (a) and 5 nm in (d).
Figure 4(a) shows isosurface renderings of the final reconstruction obtained for G40PΔN109. The reconstructed volume (4.5 nm in height) consists of six identical masses that, at threshold values accounting for 100% of the mass, do not connect with each other. Weak connections were observed at threshold values accounting for 150% of the expected mass. As opposed to the C6 architecture of G40P, the widths of the top and bottom faces of the G40PΔN109 hexamer are roughly the same. The six masses are arranged around an open channel with dimensions (4 nm in diameter) and general morphology very similar to those of the G40P oligomers.

The C6 architecture of G40P and G40PΔN109 oligomers was compared by direct superposition of the reconstructed volumes. Best fitting of the monomers of G40PΔN109 onto the monomer of G40P requires a slight upwards movement indicated by an arrow in Figure 4(b). After applying this movement to all the monomers, the G40PΔN109 volume superimposed well, and completely, within the region of the wider face of the G40P C6 structure (Figure 4(c)). Conversely, there was not a region equivalent to the narrower face of G40P in the G40PΔN109 volume (arrowhead in Figure 4(c)). From this comparison we conclude that the G40P CTD corresponds to the wider face of the C6 volume, therefore the NTD (absent in G40PΔN109) is located at the narrower face. It is then clear that, indeed, the CTD and NTD are one on top of each other along the rotational axis.

Quaternary polymorphism of G40PΔN109

In the structural studies of G40PΔN109 in buffer A presented in the previous section we did not detect any polymorphism, with all the oligomers belonging to the C6 group. However, and because polymorphism is typical in many replicative helicases,17,18,20 we wanted to explore if other buffer conditions could determine the appearance of oligomeric forms other than the C6. This task was further motivated by a recent work27 in which it is described that G40PΔN109 indeed has a quite different activity depending on the buffer composition. All these aspects considered, we analyzed the structural organization of G40PΔN109 by 2D-EM in exactly the same buffers used in the assays, the helicase activity27, that is buffers B and E. The main difference between buffers B and E is the pH value. In buffer B (pH 7.5) G40PΔN109 had almost no helicase activity, while in buffer E (pH 6.5) the mutant shows the highest helicase activity.27

The wt G40P helicase at pH 6.5 showed a quaternary polymorphism similar to the one described here (Figure 5). However, G40PΔN109 under this incubation condition displays a different behavior; it assembles into the three main different architectures described for the wt enzyme, although the C6 architecture remains as the most populated group. The C6 average images obtained for G40PΔN109 and for wt G40P at pH 6.5 share a similar organization as stated in the previous section. Similarly, the general features of the C3 average image of G40PΔN109 are comparable in shape and size to those obtained for the wt G40P. This similarity decreases when comparing the C3/C6 state of both proteins. In fact, whilst in G40P alternate monomers are clearly seen, in G40PΔN109 the six monomers are quite similar. This fact is clearly reflected in the rotational power spectra of both average images. Thus, the C3/C6 state of G40PΔN109 presents a rotational power spec-
trum in which the 3-fold component is highly reduced while the 6-fold component gets up to 90% of the energy. On the contrary, the rotational power spectrum of G40P intermediate state shows that, although the 6-fold component is the most powerful harmonic (70% energy), there is also a 3-fold component that gets up to 30% of the energy. At pH 7.5, wt G40P presents a quaternary polymorphism identical to the one described at pH 6.5, with little modifications in the percentages of each group (%) from the total number of frontal views is indicated for each 2D analysis (rows). Bars represent 5 nm.

**Figure 5.** Quaternary polymorphism of G40PΔN109. 2D-average images of G40P and G40PΔN109 oligomers after incubation in buffer E (pH 6.5) and B (pH 7.5). The pH value of these two buffers is indicated on the left-hand side of the panel. The three main different architectures (indicated at the top of the panel) were found for G40P at both pHs and for G40PΔN109 at pH 6.5. G40PΔN109 oligomers incubated at pH 7.5 were found as heptamers (C7) and as C6 group hexamers. The contour map of the symmetrized average image and the rotational power spectrum of each average image are shown. The spectrum of each average image are shown. The symmetrized average image and the rotational power spectrum of G40PΔN109 no quaternary polymorphism is found because all the hexameric particles present 6-fold rotational symmetry, and no particles with 3-fold rotational symmetry were found.

**Modelling of the NTD and CTD of G40P**

Since there is no atomic resolution information about G40P helicase, homology models were built for G40P NTD and CTD. The sequences of G40P NTD and CTD were aligned against the sequences of EcoDnaB NTD and T7gp4 CTD, respectively. Figure 6(a) shows the results of these sequence alignments, together with a secondary structure prediction for the G40P sequence. The percentages of identities and conservative substitutions between the sequences are 27% and 45% for the NTDs, and 24% and 41% for the CTDs, respectively. Moreover, there is a good agreement between the position of the predicted secondary structure elements for the sequences of the NTD and CTD of G40P and the position of the secondary structure elements in their respective homologous structures.30,31 The sequence identity and the similar distribution of the secondary structure elements strongly suggest that G40P domains share a similar fold with their homologues.

The models of G40P NTD and CTD are shown in Figure 6(b) and (c) together with the structures of their corresponding homologues. Three relevant areas (motif H4, the P-loop and an Arg-finger, Arg522 in T7gp4 sequence) are highlighted. The motif H4 (green-filled circle) has been described to interact with DNA,16 and in the X-ray structure of T7gp4 is located on the wall of the central channel. The P-loop (grey-filled circle) forms part of the NTP-binding pocket. The Arg-finger (pink-filled circle), contributed by a neighboring monomer in the case of oligomeric helicases, is also a constituent of the NTP-binding site.10,31

**Fitting of the G40P NTD and CTD into the C6 architecture**

The models of the G40P NTD and CTD were filtered to the resolution obtained for the 3D reconstructions. Figure 7(a) presents two views of the 3D model of the CTD within its corresponding low-resolution envelope (shown in yellow). Motif H4 (green), the P-loop and surrounding (grey), the Arg-finger (pink), and the N-terminal residues are indicated. The corresponding low-resolution envelope containing the 3D model of the NTD of G40P is shown in Figure 7(b).

The low-resolution models of the NTDs and CTDs were fitted into the reconstructed volume of G40P C6 architecture. The model of the CTD was manually placed within the larger mass of G40P, previously identified as the CTD. The models were positioned so that they adopt an orientation very
similar to that of the T7gp4 CTD in the crystal structure. Afterwards, the smaller and globular NTD was manually fitted into the remaining density. Under this disposition, each NTD bridges two adjacent CTDs (Figure 7(c)). The N-terminal residues of the CTD are positioned in the external face of the ring, so the more feasible monomer consists of a CTD and the NTD just below. The six CTDs orient their motif H4 towards the inside of the channel (green spheres in Figure 7(d)), the place where DNA interaction during helicase activity is assumed to take place. The region equivalent to that of Arg522 in T7gp4 (pink dots) protrudes out towards the P-loop (black dot) of a neighboring monomer, in the same way reported for the T7gp4 structure.

Fitting the G40P NTD and CTD into the 3-fold architectures

Maintaining the overall domain organization of the G40P C6 architecture (Figure 8(a)), the NTD and CTD homology models were fitted in the map of the 3-fold architectures (C3 and C3C6). In the proposed domain arrangement of the 3-fold architectures (Figure 8(b)), three of the monomers remain unmodified with respect to their organization in the C6 architecture (e.g. monomer 1 in Figure 8). The NTDs of the other three monomers (e.g. monomer 2 in Figure 8) move from their position in the C6 architecture towards the periphery of the hexamer (Figure 8(c), black arrows). This results in a new interaction between the displaced NTDs and their neighboring CTDs (Figure 8(d)). The density that connects these NTDs with their corresponding CTDs (asterisk in Figure 8(b)) would account for the hinge region, which has not been included in the hybrid model because there is not structural information (asterisk in Figure 8(d)).

The three CTDs, associated with the NTDs whose position changes, swivel upwards (Figure 8(c), red arrow) with respect to their position in the C6 architecture (Figure 8(a)). This movement results in (i) a downward movement in the inner channel of the motif H4, and (ii) the formation of two kinds of NTP-binding sites, termed type-S (short) and type-L (large). These NTP-binding sites differ by the distance between the P-loop in one monomer and the Arg-finger in the next. This separation is longest in type-L sites and shortest in type-S sites.

Discussion

The ability of G40P to adopt different quaternary organizations is shared by other members of the DnaB family (EcoDnaB), and possibly by members of different families (RepA of plasmid RSF1010 and papilloma virus helicase E1). This suggests that the helicase quaternary polymorphism may be relevant to the activity of this type of enzymes,
although no firm link has been established between polymorphism and its functionality.

Here, we have addressed the 3D characterization of G40P quaternary polymorphism. This enzyme is an adequate model to understand the initial steps of the replication process in Gram-positive bacteria. Independent 3D reconstructions of the three architectures of G40P18 have been described. Inspection of the three reconstructed volumes for G40P shows that the C3 and C3C6 architectures share a very similar organization. The principal difference consists of the alternate protrusion observed in three of the six monomers, which corresponds with the principal difference observed for the average images of top views of groups C3 and C3C6. In the EM 2D image processing, we could only detect these differences after paying specific attention to the outermost corona of the projection images. The origin of this local heterogeneity could reside in a higher flexibility on this area. Slight movements might easily produce a gradient of architectures with different values for the 3-fold and the 6-fold components in the rotational power spectra of the top views. So, the quaternary polymorphism of G40P could be viewed as an architecture where six monomers arrange with the same orientation (C6 architecture) plus several 3-fold symmetry architectures where three out of the six monomers would benefit from higher mobility than the other three (C3 and C3C6 architectures).

The characterization of G40P 3D structures in their different conformational states is certainly a first step towards a better structure–function understanding of this particular system. Additionally, the emerging view is that we are facing a situation that seems to be general to DNA replicative helicases. In this context it is natural to make comparison with previous structural data on other replicative helicases and, in particular, with EcoDnaB. Indeed, we find that this transition between the 3-fold and 6-fold states has been also observed in EcoDnaB in 2D and 3D studies17,33–35 including a model on domain assignment that is similar to the one proposed here.35 However, we provide for the first time a direct experimental evidence for the domain mapping as well as studies on the implication of some experimental conditions in the quaternary polymorphism. These findings provide a level of detail and precision that could shed light on understanding the dynamics of these ubiquitous systems.

Besides this general agreement there are, however, discrepancies among the structural studies, in
particular concerning the dimension of the reconstructions in the direction of the symmetry axis. Although this Z-compression/extension might represent a genuine structural difference between these two systems, we cannot rule out that this discrepancy is simply due to the different methodologies used for each reconstruction. The Random Conical Tilt scheme used here is a completely reference-free reconstruction method, as opposed to other approaches in which the use of a particular reference may introduce a model bias. Also, it is acknowledged that during the negative staining process in the specimen preparation for electron microscopy certain compression in the direction perpendicular to the grid may occur. If this were the case, all the images used for Random Conical Tilt reconstruction would correspond to the same structure, while reconstruction using other approaches may merge images from structures compressed in different directions.

**Function of the G40P NTD**

In an attempt to clarify the involvement of the NTD of DnaB-like replicative helicases in their quaternary polymorphism and biological activities, a mutant lacking this domain (G40PΔN109) was analyzed. Previous studies have shown that G40PΔN109 has helicase, ATPase and DNA binding activities. The results presented here show that G40PΔN109 is able to assemble as a hexamer, although the low percentages of particles, in comparison with the wt enzyme, suggests that the oligomers of the mutant protein are more unstable. Indeed, analysis by gel filtration chromatography showed that, in conditions where G40P elutes as a hexamer, G40PΔN109 elutes as a mixture of hexamers and dimers at high salt (300 mM NaCl). At low protein concentration (0.4 μM) and high salt only dimers of G40PΔN109 are found (our unpublished results). Under all conditions tested G40P elutes as a hexamer. Thus, it seems that G40P NTD is not strictly necessary for the enzymatic activities of the enzyme, although this domain could contribute to the stability of the hexamer.

**Figure 8.** Domain localization in the 3-fold architectures and the transition between symmetry states. Side and top views of the distribution of the NTD, in blue, and the CTD, in yellow, for the reconstructed C6 (a) and C3C6 (b) architectures of G40P. Numbers from 1 to 6 indicate the different monomers; N and C indicate the NTD or CTD for each monomer. The distribution proposed for the C3C6 architecture is also applicable to the C3 one. Close-up of side and top views of the domain arrangements in the C6 architecture (c) and in the 3-fold architectures (d). For clarity only three monomers are shown. The P-loop (filled black semi-sphere) together with the Arg-finger of a neighboring monomer (filled pink semi-sphere) make up an NTP-binding site, indicated by an ellipse. The motif H4 (filled green semi-sphere) is indicated. All six NTP binding sites in the C6 architecture are identical. However, two kinds of NTP-binding sites, type-L and type-S, exist in the 3-fold structures. Three alternate H4 motifs change their location on the wall of the channel in the 3-fold architectures. The asterisks in (b) and (d) indicate density linking the NTD to the CTD, proposed to account for the hinge region.
The Random Conical Tilt approach was used to calculate the non-biased initial reference volume for the refinement of the G40PΔN109 reconstruction. At threshold values accounting for 100% of the mass, no connections could be observed between the six identical masses that made up the C₆ architecture of G40PΔN109, although they must exist since functional hexamers are observed. Connectivity among these regions showed up only at low thresholds (150% of the expected mass). As stated before, G40PΔN109 hexamers seem to be structurally less stable than G40P ones. So, the regions of contact between neighboring monomers in G40PΔN109 could be less structured and more flexible. Consequently, the averaging process intrinsic to the 3D reconstruction might have reduced the averaged density in these regions.

We have shown that G40PΔN109 is able to oligomerize into the three architectures described for the wt enzyme, although the C₆ architecture was always the most predominant one. Therefore, the CTD plus the linker region of G40P are enough to display the quaternary polymorphism of the wt enzyme. A similar situation is found for the RepA helicase of plasmid RSF1010, which constitutively lacks an NTD equivalent to those of DnaB or G40P. Instead of an NTD, RepA possesses a short track of four amino acid residues (the N-terminal hook), revealed to be critical in the formation and maintenance of the hexameric structure. Under certain experimental conditions, RepA oligomers were found to exist in 6-fold and 3-fold organizations. This seems to indicate that the helicase (C-terminal) domain plus the linker or the N-terminal hook might be sufficient to initiate the structural reorganization leading to the transition between the C₆ and C₃ symmetry states.

In contrast to the observed difference in wild-type G40P, the top and bottom faces of the reconstructed G40PΔN109 volume are very similar. Interestingly, the mutant protein is able to unwind dsDNA with both 5’ to 3’ and with 3’ to 5’ polarity. It is tempting to correlate the lack of functional polarity in G40PΔN109 with a decreased structural polarity of the reconstructed volume, in contrast with the known functional polarity of the wild-type enzyme.

Summarizing, we find that the presence of the G40P NTD confers stability to the hexamer and structural and functional polarity to the enzyme, and facilitates the transition between the different architectures.

The symmetry transition

Comparison of the C₆ volumes of G40P and G40PΔN109 allowed us to locate the NTD and CTD within the C₆ architecture. This information was used to understand the domain movements that could explain the organization of the 3-fold architectures (C₃ and C₃C₆), which share a very similar 3D arrangement.

In the transition from the C₆ architecture to the 3-fold architectures, three of the monomers suffer minor modifications whereas the other three undergo significant changes. The NTDs of the later monomers move away to the periphery of the structure. This results in a modification in the structure of the hinge region, which would then exist in two different conformations, retracted and extended. Three of the hinge regions would adopt an extended conformation and the other three would be in a retracted one, similar to the one adopted by all hinges in the C₆ architecture.

The six CTDs in the C₆ architecture arrange in an identical orientation, similar to the heptameric form of T7gp4. Consequently, all the intermonomeric interfaces are equivalent. However, in the 3-fold architectures, three of the CTDs undergo a swiveling within the plane of the ring similar to the one described for the hexameric and helical form of T7gp4. As a result, the interfaces between monomers change and two different types of NTP-binding sites appear: type-S and type-L. In the type-S site a shorter distance between the P-loop and the Arg-finger is observed than in the type-L site. This correlates with biophysical studies that showed that EcoDnaB possesses two types of NTP-binding sites, with low and high affinity, and performs ATP hydrolysis in a bimodal fashion. These studies were carried out at pH 8.0, for which only the C₃ architecture has been observed.

The rotation between neighboring monomers observed in T7gp4 has been proposed to be the motion responsible for the DNA translocation. In helicases with quaternary polymorphism, the alternation between the different architectures would produce similar conformational changes needed for DNA translocation. The movements of the CTDs that swivel upwards, in the transition from the C₆ architecture to the 3-fold architectures, could be the mechanism to pull the DNA (Figure 9). A motif H4 in one monomer, interacting with DNA, descends in the transition, making possible a new interaction of the DNA with a neighboring motif H4. A new transition from the 3-fold architecture to the 6-fold (C₆) architecture returns the structure to the initial step. The DNA–protein interaction would involve one or two subunits, as was described for T7gp4.

Conclusion

Here we present for the first time 3D reconstruction of the three different conformational states of the G40P helicase. Also, an experimentally based domain assignment is performed. From these results, it is proposed that the symmetry transition is directly linked to DNA translocation. Indeed, it is important to note that Gram-positive G40P helicase would undergo similar rearrangement to transit between the symmetry states to Gram-negative EcoDnaB helicase. Although the exact mechanism for how the NTP hydrolysis energy is coupling to the DNA unwinding remains unknown, direct experimental data suggest that the simultaneous presence of the different architectures is needed for
the enzymatic activity. Thus, G40PΔDN109 shows helicase activity only under conditions where quaternary polymorphism is feasible (this work).

In addition, the biochemically inactive helicase-loader complex of *E. coli* (DnaB$\cdot$DnaC) was observed to exist only in the C3 architecture. Consequently, the existence of several architectures seems to be important in order to orchestrate the different protein–protein and protein–DNA interactions required to accomplish DNA recognition and loading onto the origin of replication. The alternation between these architectures could help in the transduction of the signal generated in the NTP-binding pocket and in the translocation and unwinding along the DNA.

**Materials and Methods**

**Electron microscopy**

G40P and G40PΔDN109 were purified as described. Samples were dialyzed against buffer A (50 mM Hepes (pH 7.0), 50 mM NaCl, 10 mM MgCl2, 1 mM dithiothreitol, 1 mM ATP) for Random Conical Tilt experiments, and against buffer B (50 mM Hepes–NaOH (pH 7.5), 1 mM DTT, 2 mM ATP, 10 mM MgCl2, 15 mM NaCl, 5% (v/v) glycerol, 50 µg/ml BSA) or E (50 mM Pipes–NaOH (pH 6.5), 1 mM DTT, 2 mM ATP, 10 mM MgCl2, 15 mM NaCl, 5% glycerol, 50 µg/ml BSA) for the analysis of the relationship between structure and helicase activity. Subsequently, diluted solutions of G40P and G40PΔDN109 (corresponding to 0.7 µM when referred to the hexamers) were adsorbed onto glow-discharged collodion/carbon coated grids, stained with 2% (w/v) uranyl acetate, and visualized in a Jeol 1200 EX-II transmission electron microscope at 60,000 magnification and 80 kV accelerating voltage. Micrographs were taken on Kodak SO-163 plates. Tilting pairs were recorded first at 50° and then at 0°, using a eucentric goniometer and low-dose conditions.

**Image processing**

Micrographs were digitized in a Zeiss-Integraph scanner (Photoscan TD model) at a pixel size of 4.1 Å. The software packages Xmipp and SPIDER was used for image processing and classification. A total number of 13,743 particles for G40P, and 4,195 for G40PΔDN109 were selected (80 × 80 pixels) for the Random Conical Tilt experiment. In the analysis of the relationship between structure and helicase activity, 7,832 and 18,013 particles were selected for G40P at pH 6.5 and 7.5, respectively, and 7,499 and 6,433 particles for G40PΔDN109 at pH 6.5 and 7.5, respectively. Particles were centered by cross-correlation against a radially symmetrised average image of the unaligned population. Rotational power spectra were then calculated between radii of 2–8 nm, covering the whole particle. A self-organizing map algorithm (KerDenSOM) was employed to classify the initial population, based on variability among the rotational power spectra. Groups having the 3 or 6 rotational harmonic as the highest one were selected. Each subset was translationally and rotationally aligned by cross-correlation methods and the so-called pyramidal system for pre-alignment construction. At this stage, the spectra were calculated only on pixels between radii of 5–8 nm. Final assessments of the homogeneity of the image subsets were carried out by KerDenSOM analysis of both, the rotational power spectra of the images, and the images themselves.

**3D reconstruction**

3D reconstructions were performed following the Random Conical Tilt scheme. Projection matching 3D orientation searches were carried out with SPIDER. For G40P, showing quaternary polymorphism, only the tilting counterpart images of the top views selected after classification were used in the reconstruction of each class. For G40PΔDN109, which did not show polymorphism, all the untitled and tilted images were used in the angular refinement. Resolution was assessed by the Fourier Shell Correlation criterion. The final 3D reconstructions were calculated with the ART algorithm. The resulting volumes were filtered to their estimated resolution. Isosurface renderings of the recon-

---

**Figure 9.** Mechanism for DNA translocation. The conformational changes in the transition from the C6 architecture (a) to the 3-fold architectures (b) would allow the translocation of DNA. A new transition to the C6 state (c) returns to the initial state but with a movement of the DNA from one subunit to another. The different states interacting with the DNA are shown from the channel. The CTDs and NTDs are represented in light and dark grey, respectively. The DNA is represented as a tube. Interactions between DNA and the motif H4 of the subunits are depicted with black asterisks.
structured electron densities were displayed using AMIRA, at a threshold value assuming a mean protein density of 1.33 g/cm³.48

Homology modeling

The G40P sequence (Q38152) was obtained from the ExPASy Molecular Biology Server. Secondary structure predictions were calculated for the whole amino acid sequence of the G40P polypeptide using the PSIPRED Protein Structures Prediction Server. Homology models were built with the SWISS-MODEL Protein  § Modelling and the CTD of T7gp4 (pdb accession code 1b79) and the CTD of T7gp4 (pdb accession code 1cr2), respectively. Homology models were filtered to 24 Å and manually fitted to the G40P reconstructions using AMIRA.

Electron Microscopy Data Bank accession numbers

The maps have been deposited at the EMBL-EBI Electron Microscopy Data Bank (EMDB). The identifier codes are 1135 for the C₆ architecture, 1146 for the C₆C₆, 1147 for the C₃, and 1148 for the G40PΔN109 mutant map.

Acknowledgements

We are very grateful to Dr C. San Martin and Dr S. H. W. Scheres for their expert advice. We thank C. Patiño, R. Arranz and C. Rollo for their help at the Electron Microscopy Facility of the Centro Nacional de Biotecnología. This work was partially supported by grants BFU2004-00217 from Dirección General de Investigación, Ministerio de Educación y Ciencia (MEC), IRO1H67465-8 from National Institutes of Health, FP6-502828 (3D-EM Network of Excellence) from European Union, G03-185 (IM3 Network) from Fondo de Investigaciones Sanitarias to J.M.C.; BIO 2001-4339 from MEC to J.M.C. and J.C.A.; BMC2003-00150 from MEC and S-0505/MAT/0283 from Comunidad Autónoma de Madrid to J.C.A.; BMC 2003-01969 from MEC to S.A. R.N. is recipient of a Pre-doctoral Fellowship from the CSIC (Programa I3P), supported by the European Social Fund. A Postdoctoral Fellowship from CAM partially supported Y.R. and P.M. L.E.D. holds a contract funded by Programa Ramón y Cajal.

References


Edited by J. Karn

(Received 23 September 2005; received in revised form 26 January 2006; accepted 27 January 2006)
Available online 9 February 2006