Internal unwinding mechanism of a replicative helicase

Taha Shahid[1,2], Muhammad Tehseen[1], Lubna Alhudhali[1], Alice Clark[1], Christos G. Savva[1], Samir M. Hamdan[1] and Alfredo De Biasio[1,2]

[1]Bioscience Program, Division of Biological and Environmental Sciences and Engineering, King Abdullah University of Science and Technology, Thuwal 23955, Saudi Arabia.

[2]Leicester Institute of Structural & Chemical Biology and Department of Molecular & Cell Biology, University of Leicester, Lancaster Rd, Leicester LE1 7HB, UK.

Correspondence should be addressed to S.M.H. (samir.hamdan@kaust.edu.sa) or A.D.B. (alfredo.debiasio@kaust.edu.sa)

Abstract

Ring-shaped hexameric helicases are nucleotide hydrolases that unwind double-stranded DNA into single strands, a necessary step in DNA replication. Critical to understanding their function are two outstanding questions: the location of DNA strand separation within the helicase, particularly at the onset of unwinding, and the precise dynamic linkage between nucleotide hydrolysis and DNA translocation. We explored these questions by employing cryo-EM to visualize the translocation mechanics of a AAA+ helicase (the SV40 Large Tumor Antigen, LTag) on forked DNA. We find the DNA fork nexus is positioned deep within the core helicase domain, with each DNA-binding loop from the six subunits securing the tracking strand by forming a paired staircase spiral which substitutes the passive strand. This structure forges an internal separation wedge,

channeling the passive strand through a gap in the subunit C-tiers at the back of the helicase. Via cryo-EM continuous heterogeneity analysis we capture, and exhaustively model, a seamless spectrum of conformations of translocating LTag with near-atomic precision. ATP hydrolysis at the tightest inter-subunit interface operates like an 'entropy switch', triggering coordinated rigid-body rotations of the subunit C-tiers and DNA-binding loops, resulting in directional escorting of the tracking strand within the central channel, concomitantly setting preparatory steps for cycle restart. Overall, we demonstrate a dynamic model for hydrolysis-coupled translocation and internal DNA unwinding by a model helicase active in replication, and potentially during initial origin melting. Structural conservation of core helicase domains suggests the mechanism may be applicable to hexameric helicases across species.

Main

The Large Tumor Antigen (LTag) of simian virus 40 (SV40) is a crucial helicase, facilitating both initiation and progression of viral DNA replication [1-3], paralleling the roles of the Cdc45–MCM–GINS (CMG) helicase in eukaryotes [4-6]. LTag bipartite architecture comprises an N-terminal origin DNA binding domain (OBD) for origin recognition and a C-terminal helicase domain with an ATPase motor for translocation [7-10]. To initiate replication, LTag monomers or dimers form head-to-head hexamers that encircle duplex DNA at two specific origin positions, causing local origin melting [11-15]. During the replication elongation phase, the two hexamers commence ATP-driven DNA unwinding, with each helicase pulling its respective tracking strand in opposite directions, thereby shearing the DNA [16]. The head-to-head configuration of the two LTag hexamers around

duplex DNA necessitates that each hexamer expels one strand from its interior, enabling the helicases to move past each other and form separate replication forks [17,18].

The structural mechanisms underlying LTag-induced origin melting, translocation and DNA unwinding remain elusive, largely based on early electron microscopy (EM) data of low resolution [11,17,19-22] or crystallographic models of LTag hexamers in the absence of DNA [8,9]. Central to unraveling these mechanisms are two pivotal questions: the process by which the LTag hexamer assembles on DNA and how the tracking strand is discriminated, as well as the nature of the allosteric communication that conveys nucleotide status across active sites to synchronize ATP hydrolysis with DNA movement.

To address these questions, we determined several cryo-EM structures of hexameric LTag bound to forked DNA. Utilizing 3D variability analysis [23], we captured a continuous spectrum of LTag translocation states which links nucleotide processing with DNA progression, allowing the first visualization of DNA coordinated escort through the helicase channel.

Structure of LTag bound to forked DNA prior to ATP hydrolysis

To begin our study of LTag translocation dynamics, we immobilized LTag on DNA in a pre-ATP hydrolysis state by reconstituting it with forked DNA and ATP, but without the Mg^{2+} cofactor, preventing nucleotide hydrolysis. Using this setup, we obtained a cryo-EM reconstruction of the complex at a resolution of 3.1 Å (Figure 1a; Extended Data Figure 1; Extended Data Table 1). The helicase core of LTag is structured as a two-tier hexameric ring. Its six oligomerization domains (N-tiers) shape a collar, with the bulkier ATPase motor domains (C-tiers) lying beneath. The OBD and J-domains exhibit

significant flexibility, as indicated by the fuzzy density in 2D class averages (Extended Data Figure 1). The forked DNA junction is nestled deep within the collar central channel, with the tracking strand held firmly by the six DNA binding loops in the C-tier chamber, while the unpaired passive strand is flexible and not visible (Figure 1b-c). The N-tiers slightly deviate from a six-fold rotational symmetry. However, a significant departure from this symmetry is observed in the C-tiers. This argument is deduced by the inspection of the "apo" configuration of LTag derived from DNA-absent particles (Extended Data Figure 2; Extended Data Table 1), which mirrors the C-tier six-fold symmetry previously identified in crystallographic analyses [8,9] (Extended Data Figure 3). The asymmetry in DNA-bound LTag arises due to the systematic histidine-nucleobase and lysine-phosphate engagements of the internal loops with the DNA tracking strand (Figure 1c). Such interactions drive progressive rotations in the C-tiers, fostering the unique staircase configuration of the loops in alignment with the DNA backbone (Extended Data Figure 3). In both DNA-anchored and apo forms, each subunit associates with an ATP molecule. However, in the apo form the internal loops appear equiplanar and flexible (Extended Data Figure 3).

At the fork junction, the cryo-EM density map shows clear definition for the DNA nucleotides over several base pairs. Beyond these five base pairs, density is more limited to the nucleobases (Figure 1b), suggesting that the DNA segment passing through the N-tier collar is partially flexible. A conformational distinction between the bound DNA segments is observed. The duplex region conforms to the typical B-DNA structure with a standard helical diameter of ~20 Å. In contrast, the tracking strand is compacted with a reduced helical diameter of ~16 Å. The side chains of the six H513 residues follow the

right-hand helical pattern of the tracking strand, establishing hydrogen bonds with their matching bases, and the tracking strand positioning is further anchored by salt bridges between K512 and R456 and the backbone phosphates (Figure 1c). Due to the mixed composition of the DNA substrate (Figure 1b), tracking and passive strand bases were assigned as thymidines and adenines, respectively. Nonetheless, the aromatic rings of pyrimidines and purines would both present a polar atom (O or N) at equivalent locations, available to interact with H513's sidechain protonated nitrogen at the **ε** (NE2) position. At the nexus, the imidazole rings of H513 from the three top loops of the staircase intrude into the DNA minor groove. At this uppermost point, the nucleobases T1 and T2 from the tracking strand not only interact with H513 but also establish Watson-Crick bonds with their complementary bases P1 and P2 on the passive strand (Figure 1c). Yet, the third base (T3) from the top may not pair with its passive strand counterpart due to interference from nearby loop residues. From this position downward, the DNA binding loops essentially occupy the space where the passive stand would be located, creating an internal separation wedge. This suggests that the passive strand, though not directly observable in the map, exits at the back through the opening between the C-tiers of the top and bottom subunits (Figure 1d). As further elaborated in this work, strand internal melting and exclusion through the C-tiers is also supported by a previously determined structure of the LTag dimer bound to origin DNA [7].

Intersubunit interactions and ATP binding

Unlike the DNA-free form of LTag, where inter-subunit interfaces appear homogenous, the DNA-bound variant in the pre-ATP hydrolysis state presents interfaces with varying features, exhibiting unique nucleotide coordination patterns corresponding to the specific

placements of the associated DNA binding loops on the staircase (Figure 2a-b). Previous crystallographic studies have revealed that the conformation of most LTag *cis*-residues (those from the same monomer) that bind to ATP is maintained when bound to ADP [8]. This observation led to the identification of critical *trans*-residues (those from the neighboring monomer) like K418 and R540, often dubbed "fingers," and R498. These residues, critical for ATP coordination and hydrolysis, serve as markers to determine the interface type. According to this classification, the two interfaces at the staircase top (A/B and B/C) in our structure conform to an ATP-like profile, characterized by the direct interactions of K418 side chain with ATP α, β and γ-phosphate, R540 with ATP γ-phosphate, and R498 with *cis*-D474, with R498 potentially interacting with the γ-phosphate via a water molecule [8] (Figure 2b). Conversely, the C/D interface closely resembles the ADP-configuration, marked by a direct interaction between R498 and the nucleotide β-phosphate. The subsequent interfaces along the staircase (D/E and E/F) align with the ADP-configuration, yet the R498 side chain is positioned at a distance that prevents direct interaction with the nucleotide. Given the staircase architecture, the bottom-most subunit (F) does not establish a pronounced interface with its top counterpart (A) (Figure 2b). The extent of the buried surface area (BSA) at the interfaces is in correlation with the nucleotide configuration (Figure 2b). Notably, ATP-like interfaces are markedly more compact than the ADP-like ones. We analyzed the nucleotide coordination in this structure and compared it to a 3.4 Å-resolution structure derived from LTag reconstituted with ADP and a forked DNA substrate with an extended tracking strand (Extended Data Figure 4; Extended Data Table 1). The general structural framework between the two is largely similar, including the arrangement of the DNA binding loops

with the DNA (Extended Data Figure 5a). Yet, beyond the missing density for the fork nexus in the ADP-bound structure—a discrepancy likely arising from the extended tracking strand resulting in varying positioning of the hexamer relative to the nexus, and tightened mask around the protein and bound DNA —there are noteworthy differences in nucleotide interaction (Extended Data Figure 5b). In the ADP-bound structure, the nucleotide coordination at the A/B and B/C interfaces is absent, due to the missing γ-phosphate in the nucleotide, the staircase positioning of the interfaces, or the partial occupancy of the nucleotide at the binding pocket (Extended Data Figure 5c). As a result, the A/B and B/C interfaces are less compact than in the ATP-bound structure (Figure 2 and Extended Data Figure 5b). On the other hand, the C/D, D/E, and E/F interfaces, remain consistent in both ATP- and ADP-bound structures. This consistency further validates the correct configuration assignment of the ATP-bound interfaces. Interestingly, in the ADP-bound map, density at the γ-phosphate position is noticeable at the C/D and D/E interfaces (Extended Data Figure 5c). This implies that the γ-phosphate might have been substituted by a chloride ion, a feature previously documented in the E1−DNA structure bound to ADP [24].

We further characterized LTag capturing its cryo-EM structure when bound to a forked DNA with the extended tracking strand, a non-hydrolysable ATP analog (AMP-PNP), and the Mg^{2+} cofactor (Extended Data Figure 6; Extended Data Table 1). While the map density at the DNA fork nexus is weaker, the structure bears a strong resemblance to its ATP-bound counterpart, including the engagement of the DNA binding loops with the DNA tracking strand (Extended Data Figure 7a-c). However, moderate rotational differences in the C-tiers and their respective DNA binding loops are observed—most

notably in the F subunit, which is inherently more flexible. Consequently, the DNA is slightly shifted compared to the ATP-bound structure (Extended Data Figure 7a). Every subunit is associated with an AMP-PNP molecule, its γ-phosphate coordinating with the Mg^{2+} cofactor (Extended Data Figure 7b). The density of AMP-PNP at the F subunit appears diminished (Extended Data Figure 7a). This could be attributed to either the inherent flexibility of the entire subunit or a reduced nucleotide occupancy. Analysis of the interfaces classifies A/B, B/C and C/D as ATP-like, and D/E and E/F as ADP-like (Extended Data Figure 7b). The interfaces generally seem tighter compared to the ATP-bound structure, especially the C/D interface which exhibits an altered nucleotide coordination (Extended Data Figure 7b). Cumulatively, when examining LTag assemblies where hydrolysis is obstructed, data suggests that ATP hydrolysis may occur amid subunits towards the staircase top, while ADP release may happen between those near its base.

Cryo-EM analysis of the continuous dynamics of LTag translocation

To gain insight into the molecular mechanisms driving LTag translocation, we assembled LTag on forked DNA with ATP and initiated hydrolysis by adding Mg^{2+}. This mixture was vitrified and analyzed using cryo-EM. Image processing yielded a consensus reconstruction at 3 Å resolution (Extended Data Figure 8; Extended Data Table 1). When compared to the ATP- and AMP-PNP-bound structures, a general increase in disorder is evident (Extended Data Figure 8), and particularly pronounced in specific regions such as subunits A and F, as well as the DNA. This suggests that the reconstruction embodies an averaged superposition of different conformational states arising from ATP hydrolysis. Density corresponding to the fork nexus appears discernible yet weak, likely due to the

averaging effect of various hexamer positions around the nexus, combined with instances of LTag translocating along DNA strands from fully unwound substrates. To characterize these overlapping states, the dataset underwent 3D variability analysis (3DVA) [23] using three variability components (Extended Data Figure 8). Five volumes spanning equally distanced points along each of the three principal components were further processed with EMReady [25], resulting in a significant improvement in map definition. Morphing between these volume series illustrates the dynamic actions of LTag as it translocates on DNA (Supplementary Videos 1-3). To elucidate the connection between ATP hydrolysis and translocation, molecular models were built into each of the 3DVA volumes by molecular dynamics flexible fitting [26]. The global atom resolvability scores (Q-scores [27]) of the refined models approach the average values reported for models built into high-resolution (~3 Å) cryo-EM maps [27] (Extended Data Table 2), enabling the detailed tracking of molecular interaction changes throughout the process.

Direct visualization of ATP hydrolysis coupled to DNA translocation

In the dynamic landscape captured by the first 3DVA component, the N-tier collar of the LTag remains nearly stationary, providing a stable base, while the C-tiers engage in rotational movements (Figure 3a). These rotations are closely synchronized with the longitudinal movement of DNA, providing an explanatory mechanism for the DNA guided linear traversal through the helicase channel (Figure 3b-d). The C-tiers, along with their DNA binding loops, rotate in a unified manner, acting as single, rigid units (Figure 3d). In the initial phase of the trajectory, we observe concerted interactions between the DNA and helicase subunits A through E (Figure 3b-c). Specifically, residues H513 and K512 engage with their respective nucleobases and phosphate groups (Figures 3b-c).

Progressing along the trajectory, there is a notable shift in interactions; the loop from subunit A disengages entirely from the DNA, while the loop of subunit B exhibits a reduced interaction (Figure 3c). In tandem, subunit F loop forms a stronger connection with the nucleotide positioned at the base of the staircase (Figure 3c). This sequence of coordinated events collectively propels the DNA forward by one base through the LTag channel (Supplementary Video 4).

The analysis of inter-subunit interface evolution provides insight into the link between ATP hydrolysis and DNA translocation (Figure 3e). Across the trajectory, the B/C interface—which starts as the most stable—shows dramatic changes, with a ~25% reduction in interface area and a ~40% decrease in contact points by the end (Figure 3e). Closer examination reveals that this interface transitions from an ATP-bound state to an ADP-bound state, marked by the complete disassociation of the finger R540 from the γ-phosphate of ATP (Figure 3e). The A/B interface, in contrast, shows only a minor decrease in its buried surface area (~9%) and reduction of only one contact point (Extended Data Figure 9a). This slight change leads to a moderate disturbance of the interface, where its ATP-type conformation becomes partially altered due to an increased gap between R540 and the γ-phosphate, along with a disruption of the R498-D474 salt bridge (Extended Data Figure 9a). The other interfaces show minimal changes and remain in a state akin to that of ADP binding (Extended Data Figure 9a). The divergence in dynamics between the A/B and B/C interfaces correlates with the rotational alignment of the subunit C-tiers: A and B rotate in concert, while B and C rotate asynchronously, leading to the full disconnection at the B/C interface (Figure 3a). Throughout the trajectory at the B/C interface, cryo-EM density consistent with the γ-phosphate of ATP is present

despite changes in the coordination of interacting residues (Extended Data Figure 9a). This could be due to either the replacement of the γ-phosphate by a chloride ion post-ATP hydrolysis, as seen in the ADP-bound LTag structure (Extended Data Figure 5), or retention of the hydrolyzed phosphate within the binding pocket through interactions with residues from the same subunit (cis). MD simulations indicate that the hydrolyzed phosphate group remains trapped in the pocket (Supplementary Video 5). Notably, a recent investigation combining cryo-EM and MD simulations identified an ADP·Pi reaction intermediate in the p97 ATPase [28], supporting that the phosphate group released following ATP hydrolysis can remain confined within the binding pocket. Importantly, when the cryo-EM maps are examined at a lower threshold, visible features corresponding to double-stranded DNA become apparent at the fork nexus (Figure 3e). This observed density pattern supports that the dynamic movements captured represent DNA unwinding occurring within the inner chamber of LTag.

The first 3D variability principal component highlights ATP hydrolysis at the LTag B/C interface as a crucial molecular event, functioning as an 'entropy switch'. Disengagement of the catalytic finger Arg540 from ATP γ-phosphate triggers a transition from a more ordered/strained state to a less ordered, more relaxed state within the subunits. The transition, which is characterized by a discernible increase in atomic B-factors (Figure 3f), particularly in subunits A, B, and F, reflects this entropic release and the consequent onset of C-tier rotations relative to the N-terminal collar domains, pivotal for DNA strand progression. Significantly, these movements take place without allosteric changes in the DNA binding loops that remain attached to the DNA throughout. The analysis indicates that energy from ATP hydrolysis does not directly drive translocation

but rather removes a translocation blockage. Following this, the second principal component illustrates the preparatory steps of reconfiguring the loop staircase, priming the system for the subsequent cycle of translocation.

Dynamics of the preparatory steps for translocation cycle restart

Similar to the first component, the onset of the trajectory of the second marks a pre-hydrolysis phase (Figure 4). Progression through this trajectory reveals three key events: firstly, subunit A release from DNA; secondly, the DNA advancement through the channel; and thirdly, loop of subunit F detachment from DNA and subsequent movement upward along the loop staircase, which occurs in tandem with the closing distance between subunits F and A (Figure 4a-c; Supplementary Video 6). The rotational axes of the C-tiers exhibit variations from those observed in the first component (Figures 3a and 4a). Notably, the axes for subunits A and B demonstrate reduced collinearity, and the F subunit rotation is in the opposite direction (Figure 4a). Interface analysis indicates disruptions at both the A/B and B/C interfaces, transitioning from ATP- to ADP-bound states. This observation might suggest dual hydrolytic events; however, we deduce that the A/B interface disruption is not a primary hydrolysis event but rather a consequence of hydrolysis at the B/C interface and the resulting rotation of the B subunit C-tier. This argument is supported by the intrinsic propensity for motion of the A/B interface due to the A subunit's limited stabilizing contacts with the F subunit, reflected in the elevated B-factors for the C-tier of the A subunit compared to the B-subunit (114 versus 88 $Å^2$). Furthermore, unlike the C subunit, the B subunit ATP-coordinating residue R498 exhibits no side chain density in the trajectory initial frame, and is not engaged in a stable salt bridge with D474 (Extended Data Figure 9b). In the crystal structure of LTag bound to

ATP [8], R498 is pivotal for positioning the "apical" water molecule critical for hydrolytic activity at the active site. This water molecule, situated at the metal ion coordination sphere apex, facilitates the nucleophilic attack on ATP γ-phosphate, thereby catalyzing its hydrolysis. The increased P-value for the A/B interface, when compared with the B/C interface (Figure 3d), further supports its inherently dynamic nature and potentially less ready state for hydrolysis. Hence, it is plausible that the A subunit, while releasing from B, could still be ATP-bound. As the cycle progresses, the former A/B interface, now in the B/C position, will achieve the necessary stability to undergo ATP hydrolysis in a subsequent translocation cycle.

The disengagement of the DNA binding loop of the F subunit from DNA and its movement up the staircase, along with the narrowing gap between the F and A subunits, signifies the preparatory phase for the next cycle of translocation. Subsequent actions, like the F subunit capture of a new nucleobase and the A subunit re-engagement with DNA to rebuild the helical staircase, likely occur too swiftly to be resolved by 3DVA. Similarly, the transition of the F subunit from ADP to ATP binding is not confirmed by the map but can be inferred from the classical rotary model of helicase function [24,29-31].

The analysis of the third 3DVA component (Supplementary Video 7) reinforces the notion that the B/C interface breakage is essential for DNA translocation (Figure 4b and 4e). Here, the DNA remains static (Figure 4b), with the B/C interface retaining its ATP-bound conformation. Meanwhile, subunit F appears to be bound to ATP at either end of the trajectory. However, A establishes a stable interaction with DNA (Figure 4b), and via ATP with B when approached by an ascending ATP bound subunit F (Figure 4e).

The integration of insights from the 3DVA analysis and the structural data of LTag in its pre-hydrolysis state (Figure 1) enables us to speculate the series of conformational changes occurring at the DNA fork nexus throughout a complete ATP hydrolysis cycle (Figure 4f, Supplementary Video 8). During this cycle, the nexus moves within the LTag inner chamber, driven by the motion of the DNA binding loops that pull on the tracking strand. Concurrently, as the loop in the F subunit moves upward, restoring the staircase structure, the nexus rotates. This rotation is triggered by the incorporation of the free nucleobase at the top of the staircase into the compact helix of the tracking strand. By contrast, the section of the tracking strand gripped by the loops translates without rotation, as previously proposed for the E1 helicase translocating on single-stranded (ss) DNA [24]. This differential movement leads to the melting of the base pair at the very junction of the fork (Figure 4f). Here, the terminal base of the passive strand is unable to engage with the counterpart on the tracking strand, which is coordinated by H513 of the B subunit.

Generality of the rotary mechanism for DNA translocation

The choreography of LTag translocation on ssDNA may well apply to other hexameric helicases believed to operate via a rotary mechanism, irrespective of the different step sizes or polarity. These include the viral E1 helicase [24], bacteriophage T7 helicase [32], bacterial DnaB [33], and archaeal MCM [34]. For eukaryotes, the unique heterohexameric structure of CMG has prompted suggestions of asymmetrical translocation models [35,36]. One model was supported by the identification of only three CMG conformers bound to ssDNA, contrary to the six expected in a symmetric rotary action [35,37], and was invoked to account for the functional asymmetry within the CMG ring [5]. However, a recent cryo-EM study has revealed two additional conformers with low occupancy [38]. Thus,

considering the conservation of residues binding to ssDNA in all MCM hexamers [34], it is likely that CMG translocates on DNA via a universal, symmetric rotary mechanism. Furthermore, our interpretation of ATP hydrolysis as an "entropy switch" could explain why disabling specific ATPase subunits in CMG does not drastically affect DNA translocation/unwinding capability [5]. For example, a Drosophila CMG mutant with alterations in the Walker A motif of the MCM6 subunit retains approximately 70-80% of the unwinding efficiency of the wild type, whereas mutations in MCM3 reduce activity to about 1-5% [5]. In the structure of human CMG bound to forked DNA, MCM6-4 and MCM3-5 are the strongest and weakest interfaces, respectively [39]. Notably, in a yeast CMG conformer with the MCM6 subunit at the top of the staircase [38], MCM6 lacks an associated nucleotide, indicating the inherent stability of the MCM6-4 interface even without ATP. Our model suggests that ATP binding rebuilds the staircase, and hydrolysis releases the translocation brake. Thus, a mutated MCM6 might ascend the staircase and re-establish the MCM6-4 interface without ATP due to its intrinsic affinity. Following one translocation cycle, the mutant MCM6-4 interface would be primed for hydrolysis, but translocation proceeds as the brake has been already removed.

A general mechanism for internal DNA unwinding

The structure of LTag bound to forked DNA and ATP, in the absence of the Mg^{2+} ion, captures a critical stage in the unwinding process, positioning the fork nexus within the helicase domain. The wedge created by the DNA binding loops and the restricted inner diameter of the helicase obstructs duplex DNA from completely traversing the C-tiers, thereby serving as the central site for DNA melting. Critically, the overlay of the B and C subunits from this structure onto the crystal structure of the LTag dimer bound to the early

palindrome (EP)-half of origin DNA [7], demonstrates a striking alignment in the region across the fork nexus and DNA binding loops (Figure 5a). Such a comparison suggests that the initiation of origin melting is contingent upon the sequential assembly of the remaining LTag subunits and ATP binding (Figure 5a). This assembly would complete the helicase loop staircase, with the D, E, and F subunits loops penetrating the DNA minor groove to encircle the tracking strand fully (Figure 5a, inset). Notably, the assembly of LTag with origin DNA in the presence of ATP has been shown to initiate melting at the EP sequence, spanning 8 base pairs [40]. These bases align with the origin DNA segment that would be laterally displaced by the loops of the D, E, and F subunits in the fully assembled LTag hexamer, providing a mechanistic correlation to the melting process (Figure 5a). It is probable that a comparable melting event happens at the AT-rich half of the origin. Here, the formation of a second hexamer was shown to cause distortions in the AT-rich area, located symmetrically from the origin center, similar to the EP-region [40]. Consequently, after melting these areas, each LTag hexamer would be in a configuration ready for ATP-driven translocation.

Utilizing 3D variability analysis, we have provided a direct visualization of the coordinated escort mechanism for translocation, a concept initially postulated by Enemark and Joshua-Tor from dual crystal structures of the E1 helicase bound to ssDNA in the presence of ADP [24,29]. Our data suggest that this mechanism can function during DNA unwinding within the LTag helicase domain (Figure 4f, Figure 5b and Supplementary Video 8).

This model is in strong agreement with biochemical evidence indicating that LTag can unwind forked substrates lacking a 5’ tail [41]. Importantly, it was determined that a

minimum of 5 nucleotides in the 3' tail is necessary for efficient unwinding [41], matching the number of nucleotides needed to engage all the loops in the staircase (Figure 1). However, internal unwinding may not always be the most efficient process: as LTag advances, the melted passive strand, exiting from the back of the C-tiers, could potentially re-anneal with the tracking strand [16]. Nevertheless, when two opposing hexamers collide at the origin, they each would pull their respective tracking strand in opposite directions [16] (Figure 5c). This action, involving multiple ATP hydrolysis events, would shear and unwind the strands between the hexamers while concurrently re-winding them beyond the DNA-binding loops, effectively preventing their re-annealing and leading to the formation of single-stranded loops extending from the C-tiers (Figure 5c). Once the unwound DNA strands between the hexamers are fully stretched, they would be expelled from their corresponding hexamers, resulting in passive strands running outside the hexamers and the simultaneous release of the tracking strand loops formed at the back of the helicase (Figure 5c). Passive strand lateral pulling and ejection are likely facilitated by the curvature observed in the longitudinal axis of the LTag double hexamer bound to origin DNA captured by EM [11]. At this stage, the two uncoupled hexamers would be properly arranged for DNA unwinding through steric exclusion [17,18]. The passive strand may be ejected through the large gap between the C-tiers of the top and bottom subunits of the staircase. However, transient gaps between interfaces in both the N-tier collar and OBD ring are also necessary for this process. The known ability of LTag to oligomerize from dimers to hexamers [13,14] supports the feasibility of these gap formations.

In eukaryotic systems, similar to the SV40 system, CMG helicases arranged head-to-head at origins lead to the formation of bidirectional replication forks [4,42]. The structure

of yeast CMG double hexamer bound to duplex DNA [43] and that pertaining to the initial phase of origin melting [44] indicate that origin melting is fundamentally driven by ATP binding, DNA binding loop staircase formation, and tracking strand engagement by the loops (Extended Data Figure 10). These factors, as proposed for the SV40 system, contribute to strand separation within the helicase inner chamber and enable each hexamer to acquire a configuration competent for ATP-driven translocation [45]. Therefore, the principles underlying origin melting appear conserved across both viral and eukaryotic domains (Extended Data Figure 10). Considering that the CMG likely translocates on DNA through a general rotary mechanism [34,38], it is plausible to suggest that DNA shearing and helicase uncoupling proposed for LTag could also be applicable to this helicase. For the CMG, the involvement of the additional factor Mcm10 is necessary for helicase uncoupling [4,46]. It has been recently suggested that Mcm10 interaction with the ssDNA exposed by the shearing process could assist in removing the passive strand from the central channel of the opposing CMG helicase [47]. Interestingly, a recent structure of human MCM double hexamer bound to origin DNA prior to CMG assembly [48] revealed DNA untwisting and distortions at the hexamer junction, which may provide a further melting site to facilitate ATP-driven DNA shearing.

Methods

Cloning, expression, and purification of LTag

Full-length SV40 large T antigen (LTag) (accession number: P03070) N-terminally double 6X histidine-tagged along with TEV cleavage site sequence was cloned into pFastBac1

plasmid by GenScript. The recombinant expression plasmid was introduced into MultiBac baculoviral DNA in DH10MultiBac™ and bacmid DNA was isolated. To prepare the baculovirus, the bacmid DNA of Ltag was transfected into Sf9 cells using FuGENE® HD (Promega) according to the manufacturer's instructions. The resulting supernatant was obtained as the P1 virus stock which was then amplified to obtain P2 virus stock. P2 virus stock was then further amplified to obtain P3 virus stock for large-scale expression. LTag was expressed by transfecting 4 L of Sf9 suspension culture at a density of 2 × 106 cells/mL with LTag P3 virus for 55-60 hrs. Cells were harvested by centrifugation at 5500×g for 10 min and re-suspended in lysis buffer [20 mM Tris (pH 8.0), 50 mM Imidazole, 500 mM KCl, 1 mM Dithiothreitol, 5% (v/v) Glycerol and EDTA-free protease inhibitor cocktail tablet/50 ml (Roche, UK)]. The cell lysate was clarified by centrifugation at 95,834×g for 1 hr at 4 °C. The supernatant was directly loaded onto a 5 ml Histrap affinity column (Cytiva) pre-equilibrated with buffer A [20 mM Tris (pH 8.0), 50 mM Imidazole, 500 mM KCl, 1 mM DTT and 10% Glycerol]. The loaded column was washed with 50 ml of buffer A containing 50 mM imidazole followed by 50 ml of Buffer A containing 100 mM imidazole to remove the non-specific binding of the protein to the column. Finally, the protein was eluted with a 20 ml gradient to 500 mM imidazole using Buffer B containing low salt [20mM Tris (pH8.0), 500 mM Imidazole, 250 mM KCl, 1mM DTT, 10 % Glycerol]. The pooled fractions of LTag were treated with TEV protease to remove the His-tag and dialyzed overnight in dialysis buffer [20 mM Tris (pH 8.0), 100 mM KCl, 1 mM DTT, and 10% glycerol]. The cleaved protein was then loaded into Histrap HP 1 ml affinity column (Cytiva) pre-equilibrated with Buffer C [20 mM Tris (pH 8), 250 mM KCl, 1 mM DTT and 10% Glycerol]. The loaded column was then washed with 10 ml of Buffer D

followed by a 10 ml linear gradient of buffer B. Flow-through fractions that contain LTag were concentrated and loaded onto a HiLoad 16/600 Superdex 200 pg (Cytiva) pre-equilibrated with gel filtration buffer [20 mM Tris (pH 8.0), 100 mM KCl, 1 mM DTT and 10% Glycerol]. Collected fractions were flash-frozen and stored at −80 °C.

DNA substrates

The forked DNA substrate for reconstitution with LTag and ATP in absence of Mg^{2+} was generated by annealing the oligonucleotides 5'-GGTGATCGTTCGCTACATGTCGTCAGGATTCCAGGCAG-3' and 5'-AGACACATGGATGTAGCGAACGATCACC-3' purchased from Eurofins. The forked DNA substrate for reconstitution with LTag and ADP, LTag and ATP in the presence of Mg^{2+} , or LTag and AMP/PNP in the presence of Mg^{2+} was adapted from SenGupta & Borowiec [49], with unpaired arms changed to poly-Ts. The constituent oligonucleotides were therefore 5'-TTCTGTGACTACCTGGACGACCGGGTTTTTTTTTTTTTTTTTTTTTTTTTTTTTT-3' and 5'-TTTTTTTTTTTTTTTTTTTTTTTTTTTTTTCCCGGTCGTCCAGGTAGTCACAGA-3'. The annealed product was purchased from Eurogentec. The gap DNA substrate for reconstitution of LTAg and ATP apo complex was created by annealing oligonucleotides 5'-AGCTATGACCATGATTACGAATTG[23ddC]-3' and 5'-TTTTTCGGAGTCGTTTCGACTCCGATTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTGCAATTCGTAATCATGGTCATAGCT-3'. LTag showed no binding to this substrate without a free 3' ssDNA tail.

Buffers

The buffer used for LTag-ATP-DNA and LTag-ADP-DNA and LTag-ATP (apo) samples was based on DNA binding assay from Li *et al.* [9] comprising 1 mM ATP, 50 mM Tris-HCl (pH 8.0) and 200 mM NaCl. The buffer used in the LTag-ATP-Mg^{2+}-DNA sample was based on DNA helicase assay from [9], composed of 50 mM HEPES (pH 7.5), 1 mM ATP, 3 mM $MgCl_2$, 1 mM DTT and 50 mM NaCl. The buffer used for the LTag-AMPPNP-Mg^{2+}-DNA complex was 50 mM Tris-HCl, pH 7.5, 1 mM DTT, 100 mM NaCl, 5 mM MgCl and 3 mM AMP-PMP.

Cryo-EM sample preparation

The LTag stock was pre-incubated with buffer before addition of DNA substrate (in equimolar ratio to putative hexamer). The steps were performed in a 4 °C temperature-controlled room. The final protein concentration in samples applied to the grids was ~0.098 mg/ml (equivalent to 0.2 μM hexamer). R1.2/1.3 or R2/2 UltrAuFoil® grids (Quantifoil Micro Tools GmbH) were glow-discharged with 40 mA for 5 min inside a Quorum GloQube® Plus unit. Graphene oxide (Merck KGaA) coating was applied[50]. 3 μl of sample was applied to the grid in a VitroBot Mark IV at 4 °C and 100% relative humidity, and following a wait time of 30 s, blotted for 3s with blot-force 4, and plunged into liquid ethane. For the LTag-AMPPNP-Mg^{2+}-DNA complex, 2 μM LTag (with respect to the hexamer) was mixed in the corresponding buffer and incubated at 4 °C for 30 min. Following this, 2 μM DNA was added and the reaction incubated at 4 °C for a further 30 min. 3 uL of sample was deposited onto a R2/2 Au 300 mesh grid previously glow discharged at 40 mA for 60 sec. The grid was frozen on a Vitrobot with a 3 sec blotting

time, 0 sec wait time and a blot force of 5, with 4 degrees C and 100% humidity in the chamber.

Cryo-EM data acquisition

Data for the LTag-AMPPNP-Mg^{2+}-DNA structure were collected on a Titan Krios G4 TEM (Thermo Fisher Scientific, Eindhoven) located at the Imaging and Characterization Core Lab of the King Abdullah University of Science and Technology, Saudi Arabia. Movies in EER format were recorded at 130,000x magnification (calibrated pixel size of 0.93 Å px^{-1}) using a Falcon 4i direct detection camera at the end of a Selectris-X energy filter with a 10eV slit width. Data were collected with a defocus range of -2.5 to -1.5 μm and an accumulated dose of 38.7 e-/$Å^2$. All other data were collected on a Titan Krios G3 TEM (Thermo Fisher Scientific, Eindhoven) at 105,000× magnification in AFIS mode with 100 mm objective aperture, using the GATAN K3 direct detection camera at the end of a GATAN BioQuantum energy filter (20 eV slit width) in counted super-resolution mode (binning 2) at the Leicester Institute of Structural and Chemical Biology, UK. Individual exposures were acquired at a dose rate of 18 e^- px^{-1} s^{-1} in 50 equal fractions over 2 s, with a calibrated pixel size of 0.835 Å px^{-1} and an accumulated dose of 50 e^- $Å^{-2}$. The nominal defocus was varied between -1.0 μm and -2.5 μm in 0.3 μm steps. The gross number of movies collected from each sample were: 8,072 (LTag-ATP-DNA); 9,737 (LTag-ADP-DNA); 6,127 (LTag-AMPPNP/Mg^{2+}-DNA); and 8,460 (LTag-ATP/Mg^{2+}-DNA).

Cryo-EM Data processing

Movie stacks were corrected for beam-induced motion using RELION's implementation[51] of MotionCor2[52] with dose-weighting, and CTF estimation of the integrated micrographs was performed using CTFFIND 4.1[53]. The micrographs were denoised prior to picking by

Topaz[54], both using default models. The picks were extracted using a box size of 384, and subject to several iterations of 2D classification and selection using VDAM in RELION 4[51] (with decreasing binning factors of 4×, 2× and 1×). In general, particle subsets from above were imported into cryoSPARC (Structura Biotechnology Inc., Toronto, Canada) and re-extracted from Patch Motion Correction and Patch CTF Estimation jobs, and initially subjected to *ab-initio* reconstruction with 3 classes (allowing for DNA-bound, apo and junk categories). Where two junk classes were detected, the process was repeated using 2 classes. The resulting particle selections were further corrected for beam-induced motion with Local Motion Correction. Eventual final reconstructions were derived using non-uniform refinement, with CTF and per-particle scale refinements enabled. The LTag-ATP-DNA reconstruction was initially refined to 2.9 Å in RELION 4 using SIDESPLITTER. The final particle subset following 2D classification was subject to a 3-class 3D (VDAM) initial model job. This produced two junk classes; the non-junk class was subjected to 2-class 3D classification. This yielded apo (54%) and DNA-bound (44%) classes. The latter was iteratively 3D auto-refined employing CTF refinements and Bayesian polishing. The refined particle stack with its assignments was imported into cryoSPARC, and re-refined using non-uniform refinement. This gave a better regularised map despite a slightly lower reported global GSFSC resolution. 3D variability analysis[23] of the LTag-ATP/Mg^{2+}-DNA consensus reconstruction was carried out using its refinement mask (created as soft mask around protein and bound DNA); with 3 modes and filter resolution of 5 Å (the global GSFSC resolution being 3 Å), as well as a high-pass resolution of 150 Å. Latent coordinates exhibited a spherical/3D Gaussian distribution, characteristic of continuous conformational variability. 20 volumes (frame_000 to frame_019) were sampled along

each principal component (0, 1, 2). Data had been exhaustively cleaned via repeat classification without separating conformations. The LTag-AMPPNP-Mg^{2+}-DNA structure was wholly refined in RELION 4. Briefly, particles picked by Topaz (2,859,203 particles) were extracted with a box size of 256 pixels and 4X downscaling (3.72 Å/px), and extensively cleaned by 2D classification. The resulting particles (957,358 particles) were subjected to 3D classification with 3 classes. Particles from the best class (235,707 particles) were re-extracted to full size (0.93 Å/px) and subjected to 3D refinement, CTF-refinement and polishing, resulting in a 3.2 Å resolution map. LTag-ATP (apo) reconstruction from the gapped DNA sample was refined in RELION. Steps up to and including 2D classification and subset selection were done in RELION 4. Ab-initio reconstruction was obtained from cryoSPARC. 3D classification and auto-refinement was carried out in RELION 5 using Blush regularisation (with CTF refinement). Bayesian polishing was not performed. Total number of particles incorporated into the final reconstructions were: 92,330 (apo LTag-ATP); 210,910 (LTag-ATP-DNA); 97,587 (LTag-ADP-DNA); 235,707 (LTag-AMPPNP/Mg^{2+}-DNA); and 201,416 (LTag-ATP/Mg^{2+}-DNA). Local resolution maps were calculated in Phenix, and rendered on the full map (pre post-processing) with adjusted colouring. 3DFSC[55] metrics were obtained from web server https://3dfsc.salk.edu/.

Model building and refinement

The resulting volumes were post-processed with EMReady for purposes of interactive model building in ISOLDE[26], alongside masked full/globally sharpened maps using automatically calculated B-factors. RELION or cryoSPARC post-processed maps were also used to derive draft models from ModelAngelo to guide the subsequent process. No

form of post-processed maps were used for iterative map refinement itself. *Modelling of the LTag-ATP-DNA complex:* Initially, the crystal structure of ATP-bound LTag (1SVM) [8] was real space refined into the LTag-ATP-DNA map using PHENIX, and oligo-dT_6 from E1-ADP-Mg^{2+}ssDNA crystal structure (2GXA)[24] fitted into density in ChimeraX. The map/model association was first globally simulated in ISOLDE, followed by multiple cycles of local rebuilding and Phenix real space refine with ISOLDE's recommended settings. In latter rounds, additional bases in the tracking strand, and complementary strand were added and re-modelled. Several iterations of LocScale[56] and ISOLDE refinement were used to improve the DNA nexus map and model. *Modelling of the LTag-ADP-DNA complex:* The crystal structure of ADP-bound LTag (1SVL)[8] was flexibly fitted into the LTag-ADP-DNA reconstruction using iMODFIT and real space refined, with ssDNA incorporated as before. The model was then rebuilt and refined as previously. *Modelling of the LTag-AMPPNP-Mg^{2+}-DNA complex:* A refined model of LTag-ATP-DNA (omitting the duplex portion of DNA) was initially rigid body fitted into the LTag-AMPPNP-Mg^{2+}-DNA reconstruction in ChimeraX. Adaptive distance restraints were applied to individual chains in ISOLDE, and simulated against the map. Restraints were progressively released and model fully rebuilt. Towards the end, ATP was replaced with AMPPNP and Mg^{2+}, and resimulated/rebuilt. *Modelling of the LTag-ATP/Mg^{2+}-DNA complex into the 3DVA volumes:* The LTag-ATP-DNA model, following rigid body fitting into LTag-ATP/Mg^{2+}-DNA consensus map, was first rebuilt into a working consensus model. This model was flexibly fitted into the central-most sampled volume (frame_010) from the first principal component. The top-ranked AlphaFold2[57] monomer structure prediction for SV40 LTag was superimposed to each of the six subunits using MatchMaker[58]. Residues

not pertaining to the core helicase 265-627 were deleted. Adaptive distance restraints were applied to individual chains, followed by a whole map-model simulation in ISOLDE. Restraints were progressively released, and the model rebuilt end-to-end. The derived model was placed into two intermediate volumes (frame_005 & frame_015). Adaptive restraints were reapplied and models entirely rebuilt as above. These models were placed into terminal volumes (frame_000 & frame_019), and rebuilt as per the same procedure. The central model from principal component 0 was placed into central volume of component 1. After remodelling, it was used as basis for modelling the two intermediate, and then terminal frames as before. The process was repeated for principal component 2. Each residue in all of the above models was manually inspected/rebuilt at least once. Q-scores were calculated in Model-map Q-score[27] ChimeraX plugin using default/recommended sigma = 0.6. The consensus LTag-ATP/Mg^{2+}-DNA structure couldn't be definitively modelled throughout due to representing multiple averaged states, and was used chiefly as an interim model to generate initial model drafts for variability volumes. Modelling of the LTag-ATP (apo) complex: LTAg-ATP (1SVM) crystal structure was first Phenix real-space-refined into the apo LTAg-ATP reconstruction postprocessed with EMReady. AlphaFold2 monomer prediction corresponding to amino acids 266 - 627 was superposed with each of the six chains using MatchMaker, and subjected to molecular dynamics bulk flexible fitting with adaptive distance restraints in ISOLDE. Restraints were released inside a whole map/model simulation, followed by localised rebuilding in ISOLDE. The final Phenix real space refined structures were assigned secondary structure annotations using DSSP [59].

Map and model visualization

Maps were visualized in UCSF Chimera[60] and ChimeraX[61] and all model illustrations and morphs were prepared using ChimeraX or PyMOL.

Statistics and reproducibility

No statistical methods were used to predetermine sample size. The experiments were not randomized, and investigators were not blinded to allocation during experiments and outcome assessment.

Data availability

The data that support this study are available from the corresponding authors upon request. All cryo-EM maps and models of the LTag complexes presented in this work will be deposited in the EMBD and Protein Data Bank.

Contributions

T.S., S.M.H. and A.D.B conceived the study. M.T. purified the protein. T.S., A.C., C.G.S. and A.D.B. performed cryo-EM imaging and processing. T.S. implemented the 3DVA analysis. T.S. and A.D.B. performed model building and validation. S.M.H. and A.D.B. supervised the study. A.D.B. wrote the manuscript with contribution from the other authors.

Figures

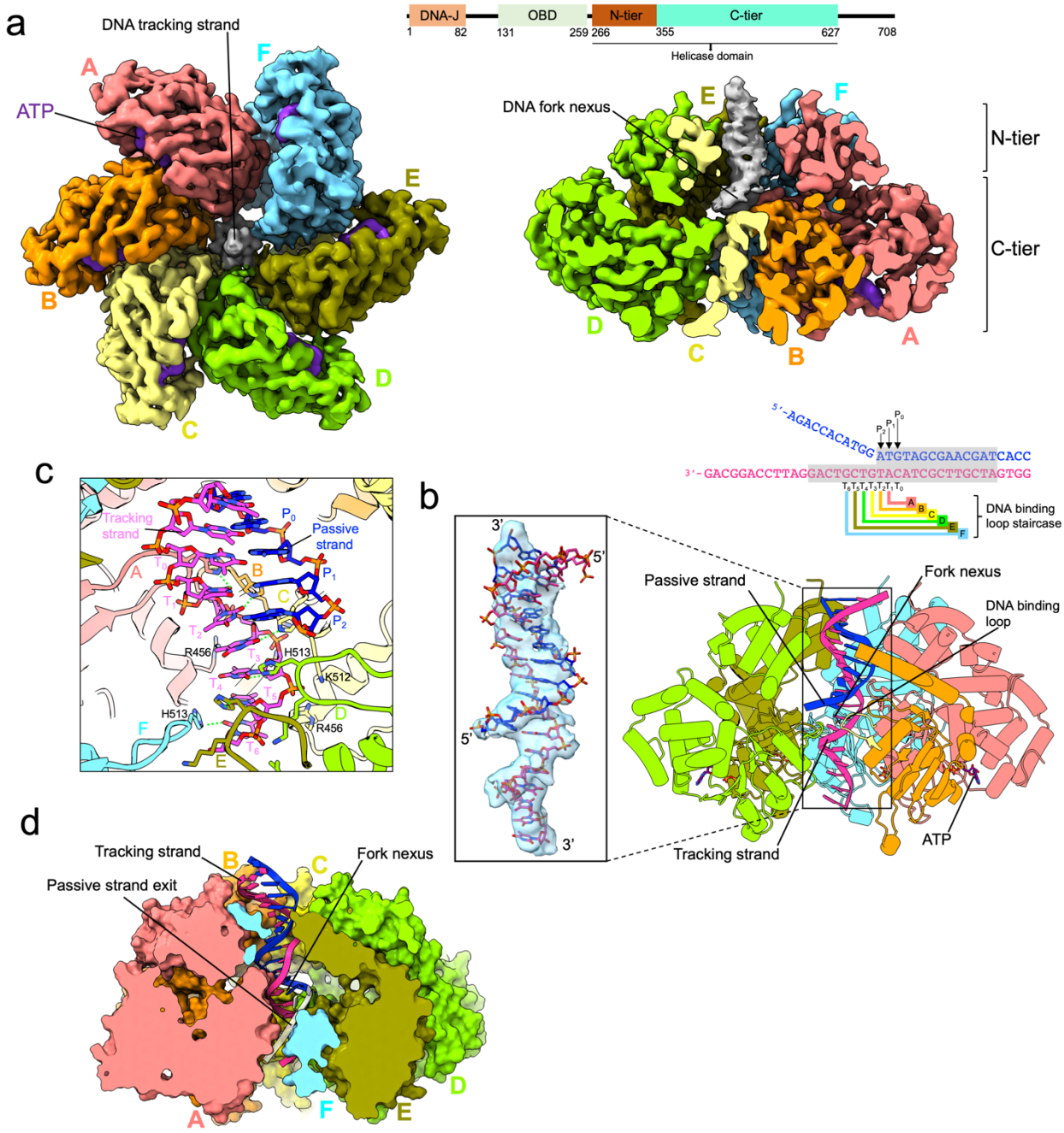


Figure 1 Cryo-EM structure of the LTag hexamer bound to forked DNA and ATP in a pre-hydrolysis state.

a Bottom and side views of the cryo-EM reconstruction of the LTag hexamer bound to forked DNA and ATP. In the side view, the map has been sliced to reveal the DNA density crossing the N-tier. Above, the domain structure of the LTag monomer is illustrated. **b** Molecular model in ribbon representation. The model has been sliced to highlight the DNA binding loops staircase intruding into the DNA minor groove at the fork nexus. The inset shows the map density around the DNA model in stick representation. **c** Molecular interactions at the DNA nexus. Hydrogen bonds between H513 and nucleobases are shown as green dotted lines. Below the P2 base of the tracking strand, the DNA binding loops of the D, E, and F subunits create an internal DNA separation wedge. **d** LTag model depicted with a surface representation that has been sectioned to show its internal structure. The tracking and passive strands are displayed as sticks in pink and blue, respectively. The 5' tail of the passive strand, which is not visible in the map, has been theoretically modeled to pass through the gap between the C-tiers of the A and F subunits and is highlighted in white.

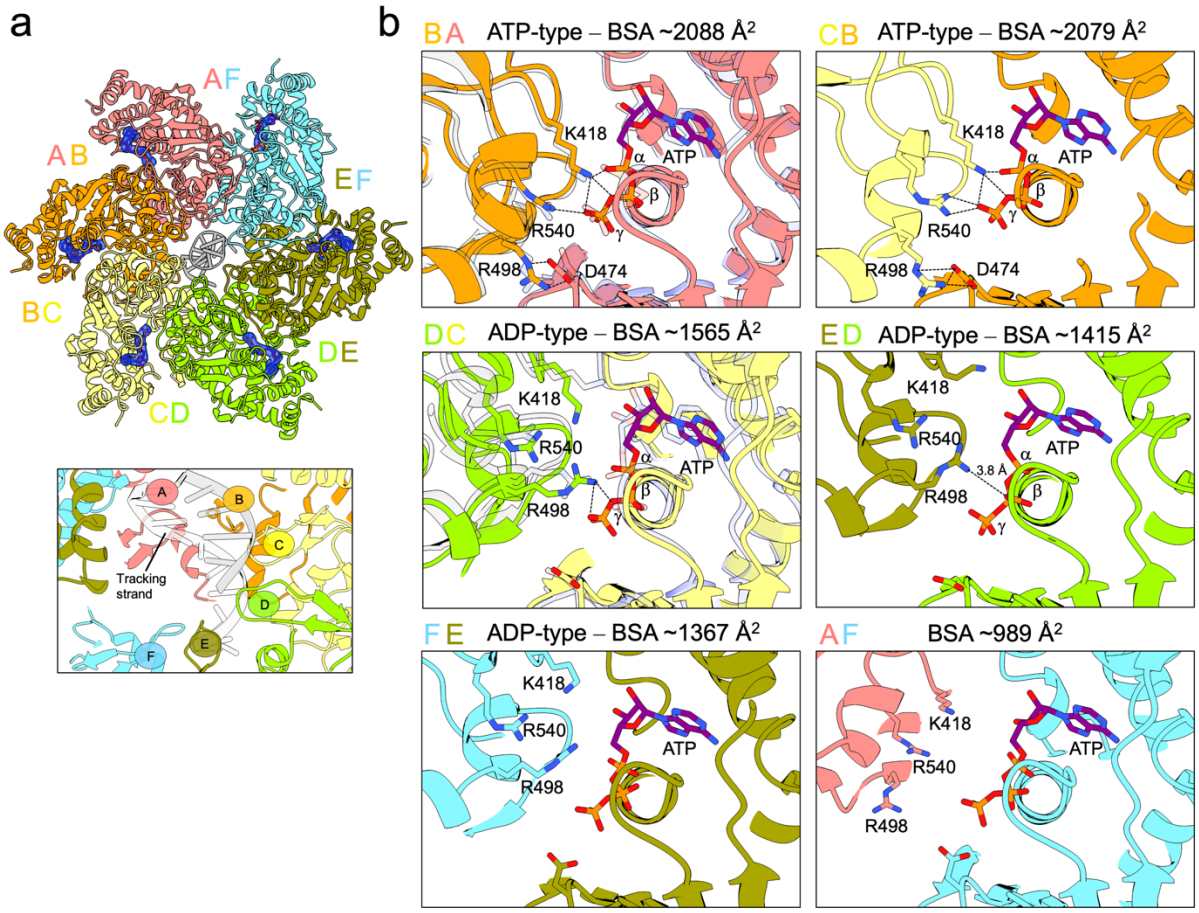


Figure 2 Inter-subunit interactions and ATP binding in the LTag–forked DNA complex

a Bottom view of the cryo-EM model. The map density around the ATP molecules is shown in blue. The lower inset shows details of the DNA binding loop staircase and DNA. **b** Inter-subunit interactions at the nucleotide binding pocket. Calculated Buried Surface Areas (BSA) for each interface are shown, along with the assigned interface type. Hydrogen bonds are shown as black dotted lines. In the insets, chain C from the LTag-ATP crystal structure (PDB: 1SVM)[8] is superimposed on chain A of the cryo-EM model at the AB interface, while chain B from the LTag-ADP crystal structure (PDB: 1SVL)[8]

overlays chain C at the CD interface, both shown in transparency to underscore the structural alignment.

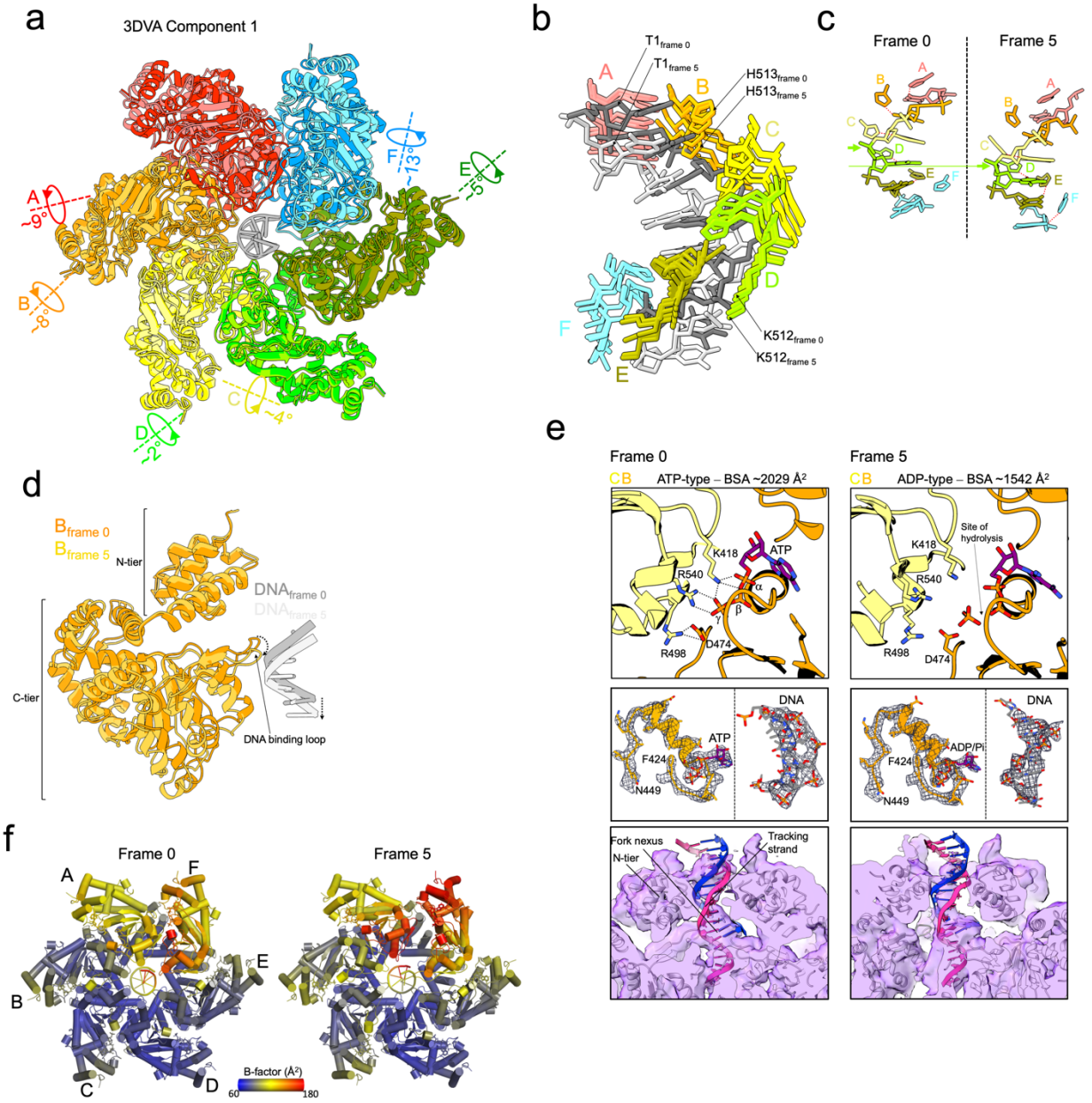


Figure 3 Visualization of ATP hydrolysis coupled to DNA translocation from the 3DVA analysis

a Overlay of the LTag models from the initial and final frames of the trajectory derived from the first 3DVA component, illustrating the extent and orientation of the C-tiers rotations in each subunit. **b** Overlay of the DNA models from the initial and final frames of the first 3DVA component, along with the superimposed models of DNA-coordinating residues K512 and H513 across the five frames. **c** The initial and final frame models of DNA and the H513 side chain from the first 3DVA component are displayed side by side. Nucleotides are color-coded to match the DNA binding loop they associate with. Hydrogen bonds between H513 and the nucleobases are depicted with dotted red lines. The green arrows indicate the phosphate group of the nucleotide engaged with the DNA binding loop of subunit D, emphasizing a net translocation of one base along the trajectory. **d** The B subunit models from the initial and final frames of the first 3DVA component trajectory are superimposed to show the C-tier rotation, which causes a downward shift of both the DNA binding loop and the DNA it binds. **e** *Top insets*: Interactions at the nucleotide binding pocket between the B and C subunits in the initial and final frames of the first 3DVA component. Calculated Buried Surface Areas (BSA) are shown, along with the assigned interface type. Hydrogen bonds are shown as black dotted lines. The analysis reveals a disruption at the interface between subunits B and C, indicative of a hydrolytic event. The location where ATP hydrolysis is proposed to occur is highlighted. *Middle insets*: Focus on map density around the nucleotide binding pocket and DNA regions in the first and final frames of the first 3DVA components, demonstrating the 3DVA maps' quality. *Bottom insets*: Cryo-EM map and model of the first and final frames of the first 3DVA component. The map is displayed at a low threshold to reveal the density throughout the N-tier channel, suggesting the presence of double-stranded

DNA. **f** Models of LTag from the initial and final frames of the first 3DVA component colored according to B-factors, highlighting an increase in B-factors in the final frame, especially within the A, B, and F subunits.

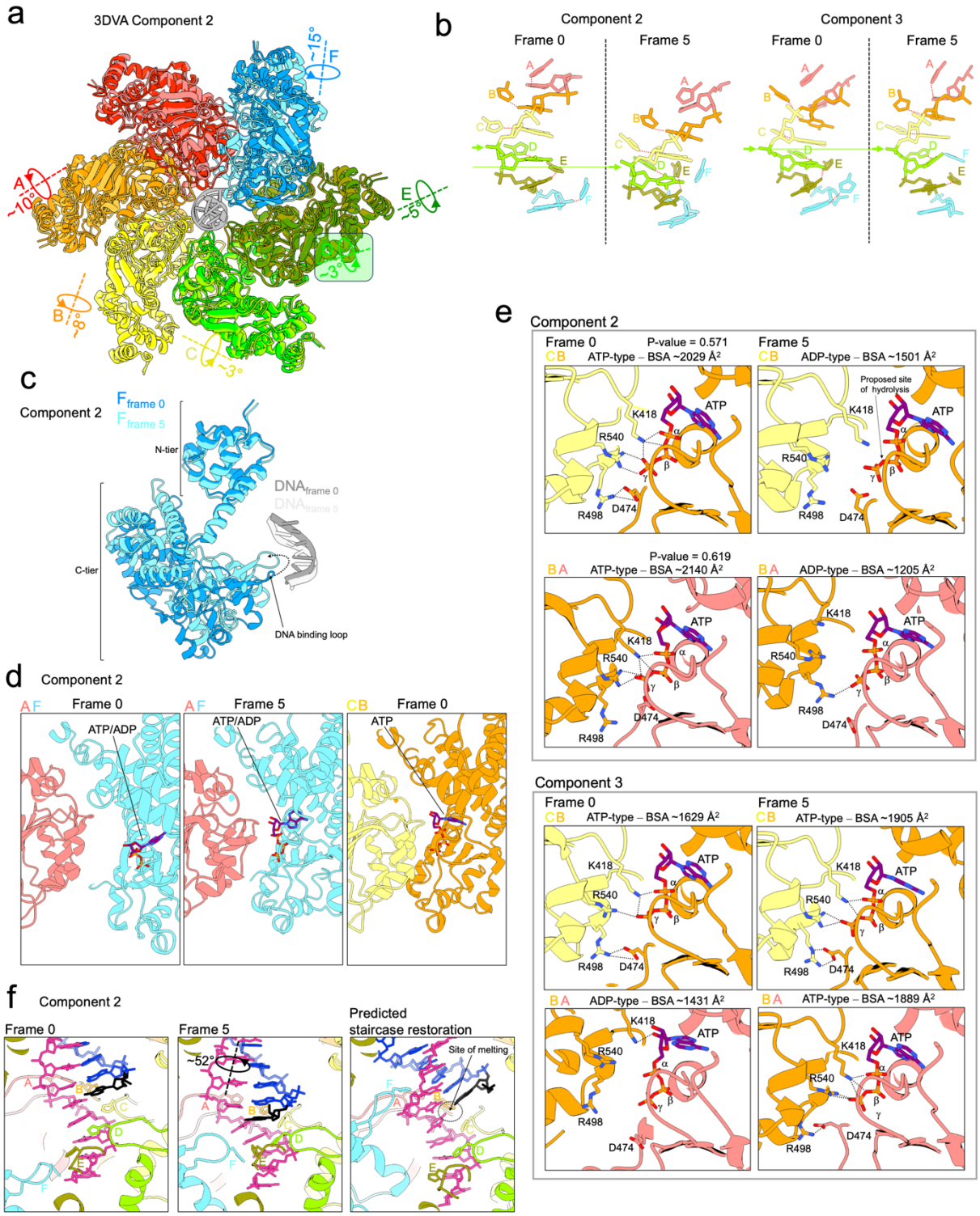


Figure 4 Dynamics of the preparatory steps for translocation cycle restart from the 3DVA analysis

a Overlay of LTag models from the initial and final frames of the trajectory, as derived from the second 3DVA component, illustrating the magnitude and direction of the C-tiers rotations across each subunit. **b** The initial and final frame models of DNA and the H513 side chain from the second and third 3DVA components displayed side by side. Hydrogen bonds between H513 and the nucleobases are depicted with dotted red lines. The green arrows indicate the phosphate group of the nucleotide engaged with the DNA binding loop of subunit D, which show a net translocation of one base along the trajectory of the second component and no translocation for the third component. **c** Overlay of the F subunit models from the initial and final frames of the second 3DVA component trajectory, showing the C-tier rotation and upward shift of the DNA binding loop, which results in the detachment of the DNA from the subunit. **d** Interfaces between the C-tiers of the A and F subunits in the initial and final frames of the second 3DVA component. A third inset presents the interface between the C-tiers of the B and C subunits in the initial frame, emphasizing the narrowing gap between the A and F subunits throughout the trajectory. **e** Interactions at the nucleotide binding pocket between the B and C subunits or A and B subunits in the initial and final frames of the second and third 3DVA components. Calculated Buried Surface Areas (BSA) or P-values are shown, along with the assigned interface type. Hydrogen bonds are shown as black dotted lines. The analysis reveals a disruption at both BC and AB interfaces in the second component, and no disruption of the BC interface in the third component. These findings collectively indicate that disruption of the BC interface is necessary for DNA translocation. **f** Prediction of the conformational

changes at the DNA fork nexus in a full translocation cycle. The duplex DNA segment is modelled based on the pre-hydrolysis structure of LTag (Figure 1). The structure in the third panel is equivalent to the initial frame of the second 3DVA component, but with the LTag subunits rotated one step around the six-fold symmetry axis of the N-tier. The DNA nexus rotation required to re-establish the loops staircase leads to the unpairing of the terminal base of the passive strand (colored in black).

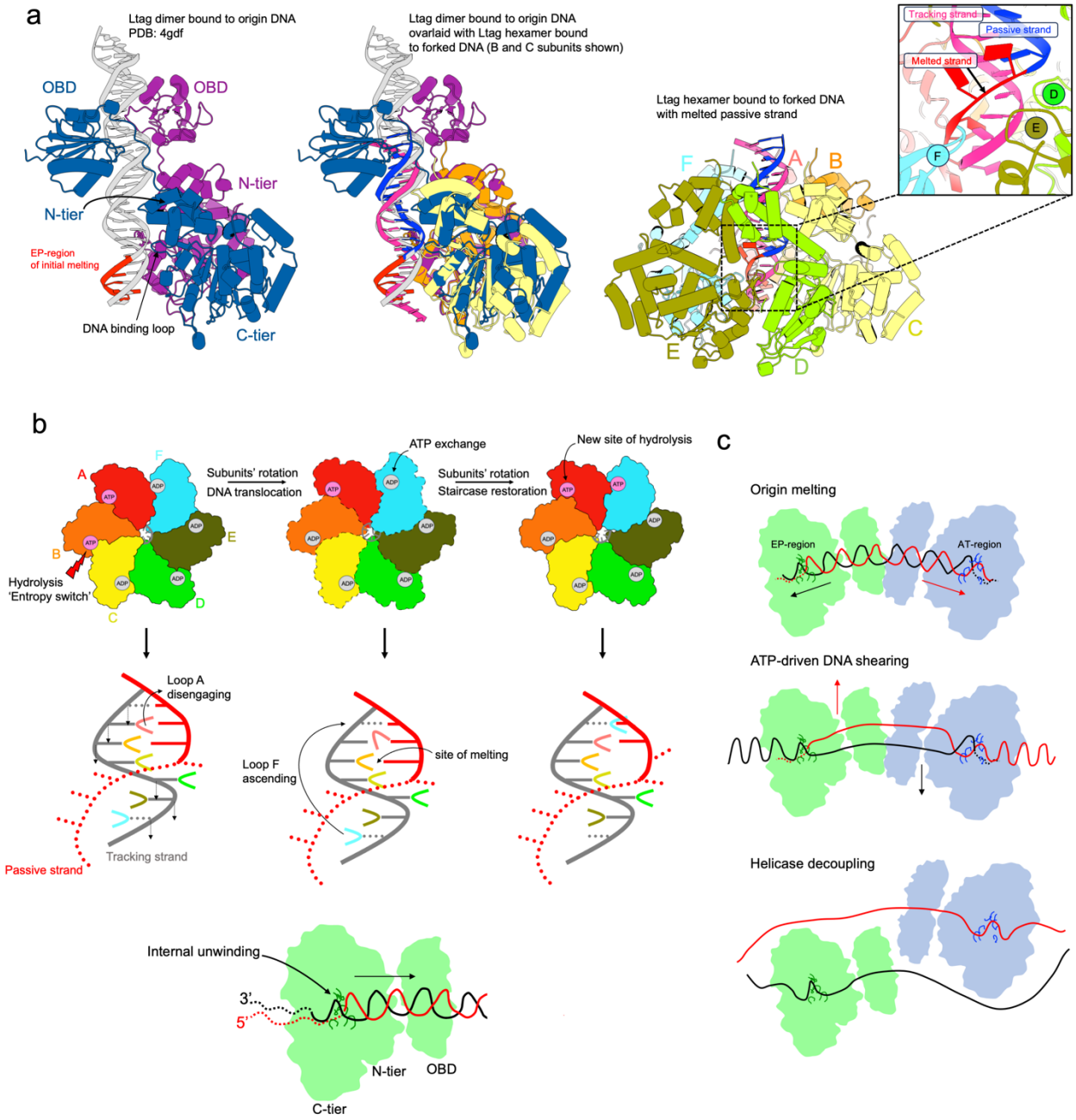


Figure 5 Model for DNA origin melting and internal unwinding by LTag

a *Left panel*: Crystal structure of Ltag dimer bound to origin DNA [7]. The early-palindrome (EP) region of origin DNA that undergoes initial melting upon LTag double hexamer assembly [40] is colored red. *Middle panel*: Superposition of the structure showed in the left panel and the cryo-EM structure of LTag bound to forked DNA determined in this work.

Only the B and C subunits of the latter structure are shown. *Right panel*: The cryo-EM structure of LTag bound to forked DNA, with the passive strand extended by 6 bases colored red, models the DNA origin melting process. It demonstrates how the intrusion of loops from the D, E, and F subunits into the DNA minor groove to fully engage the tracking strand leads to the unpairing of DNA bases. **b** Mechanism of DNA internal unwinding by LTag, synthesizing the insights from the static structure of LTag bound to forked DNA and ATP with the 3D variability analysis. This model outlines a single translocation cycle where ATP hydrolysis at the hexamer tightest interface (BC) acts as an 'entropy switch'. Hydrolysis initiates synchronized rotations of the C-tiers across different subunits and rearrangements of the DNA binding loops, facilitating the longitudinal movement of the tracking strand and DNA nexus within LTag inner chamber. Following the exchange from ADP to ATP by the F subunit, a new series of C-tier rotations occurs. This shifts the DNA binding loop of the F subunit to the top of the staircase, concurrently establishing the ATP interface between the F and A subunits. The reconstitution of the staircase structure leads to the melting of one base pair at the nexus, progressing the unwinding process. **c** Model of origin unwinding facilitated by dual LTag helicases: assembly of two hexamers in a head-to-head configuration at the origin causes melting of the AP- and AT-rich regions. The ATP-driven traction on the tracking strand by each hexamer leads to shearing of the DNA between the hexamers and the creation of single-stranded loops at the helicase domain rear. When the unwound DNA strands spanning the hexamers are entirely elongated, they exit their respective hexamers. This results in the passive strands being routed around the exterior of the hexamers while concurrently releasing the tracking strand loops formed at the helicase back, culminating in the separation of the helicases.

References

1 Fanning, E. Simian virus 40 large T antigen: the puzzle, the pieces, and the emerging picture. *J Virol* **66**, 1289-1293 (1992). https://doi.org/10.1128/JVI.66.3.1289-1293.1992

2 Stillman, B. Smart machines at the DNA replication fork. *Cell* **78**, 725-728 (1994). https://doi.org/10.1016/s0092-8674(94)90362-x

3 Waga, S. & Stillman, B. The DNA replication fork in eukaryotic cells. *Annu Rev Biochem* **67**, 721-751 (1998). https://doi.org/10.1146/annurev.biochem.67.1.721

4 Douglas, M. E., Ali, F. A., Costa, A. & Diffley, J. F. X. The mechanism of eukaryotic CMG helicase activation. *Nature* **555**, 265-268 (2018). https://doi.org/10.1038/nature25787

5 Ilves, I., Petojevic, T., Pesavento, J. J. & Botchan, M. R. Activation of the MCM2-7 helicase by association with Cdc45 and GINS proteins. *Mol Cell* **37**, 247-258 (2010). https://doi.org/10.1016/j.molcel.2009.12.030

6 Moyer, S. E., Lewis, P. W. & Botchan, M. R. Isolation of the Cdc45/Mcm2-7/GINS (CMG) complex, a candidate for the eukaryotic DNA replication fork helicase. *Proc Natl Acad Sci U S A* **103**, 10236-10241 (2006). https://doi.org/10.1073/pnas.0602400103

7 Chang, Y. P. *et al.* Mechanism of origin DNA recognition and assembly of an initiator-helicase complex by SV40 large tumor antigen. *Cell Rep* **3**, 1117-1127 (2013). https://doi.org/10.1016/j.celrep.2013.03.002

8 Gai, D., Zhao, R., Li, D., Finkielstein, C. V. & Chen, X. S. Mechanisms of conformational change for a replicative hexameric helicase of SV40 large tumor antigen. *Cell* **119**, 47-60 (2004). https://doi.org/10.1016/j.cell.2004.09.017

9 Li, D. *et al.* Structure of the replicative helicase of the oncoprotein SV40 large tumour antigen. *Nature* **423**, 512-518 (2003). https://doi.org/10.1038/nature01691

10 Luo, X., Sanford, D. G., Bullock, P. A. & Bachovchin, W. W. Solution structure of the origin DNA-binding domain of SV40 T-antigen. *Nat Struct Biol* **3**, 1034-1039 (1996). https://doi.org/10.1038/nsb1296-1034

11 Cuesta, I. *et al.* Conformational rearrangements of SV40 large T antigen during early replication events. *J Mol Biol* **397**, 1276-1286 (2010). https://doi.org/10.1016/j.jmb.2010.02.042

12 Dean, F. B., Borowiec, J. A., Eki, T. & Hurwitz, J. The simian virus 40 T antigen double hexamer assembles around the DNA at the replication origin. *J Biol Chem* **267**, 14129-14137 (1992).

13 Mastrangelo, I. A. *et al.* ATP-dependent assembly of double hexamers of SV40 T antigen at the viral origin of DNA replication. *Nature* **338**, 658-662 (1989). https://doi.org/10.1038/338658a0

14 Parsons, R. E. *et al.* Cooperative assembly of simian virus 40 T-antigen hexamers on functional halves of the replication origin. *J Virol* **65**, 2798-2806 (1991). https://doi.org/10.1128/JVI.65.6.2798-2806.1991

15 Tjian, R. The binding site on SV40 DNA for a T antigen-related protein. *Cell* **13**, 165-179 (1978). https://doi.org/10.1016/0092-8674(78)90147-2

16 Langston, L. D., Yuan, Z., Georgescu, R., Li, H. & O'Donnell, M. E. SV40 T-antigen uses a DNA shearing mechanism to initiate origin unwinding. *Proc Natl Acad Sci U S A* **119**, e2216240119 (2022). https://doi.org/10.1073/pnas.2216240119

17 Dodson, M., Dean, F. B., Bullock, P., Echols, H. & Hurwitz, J. Unwinding of duplex DNA from the SV40 origin of replication by T antigen. *Science* **238**, 964-967 (1987). https://doi.org/10.1126/science.2823389
18 Yardimci, H. *et al.* Bypass of a protein barrier by a replicative DNA helicase. *Nature* **492**, 205-209 (2012). https://doi.org/10.1038/nature11730
19 Gomez-Lorenzo, M. G. *et al.* Large T antigen on the simian virus 40 origin of replication: a 3D snapshot prior to DNA replication. *EMBO J* **22**, 6205-6213 (2003). https://doi.org/10.1093/emboj/cdg612
20 Valle, M., Chen, X. S., Donate, L. E., Fanning, E. & Carazo, J. M. Structural basis for the cooperative assembly of large T antigen on the origin of replication. *J Mol Biol* **357**, 1295-1305 (2006). https://doi.org/10.1016/j.jmb.2006.01.021
21 Valle, M., Gruss, C., Halmer, L., Carazo, J. M. & Donate, L. E. Large T-antigen double hexamers imaged at the simian virus 40 origin of replication. *Mol Cell Biol* **20**, 34-41 (2000). https://doi.org/10.1128/MCB.20.1.34-41.2000
22 Wessel, R., Schweizer, J. & Stahl, H. Simian virus 40 T-antigen DNA helicase is a hexamer which forms a binary complex during bidirectional unwinding from the viral origin of DNA replication. *J Virol* **66**, 804-815 (1992). https://doi.org/10.1128/JVI.66.2.804-815.1992
23 Punjani, A. & Fleet, D. J. 3D variability analysis: Resolving continuous flexibility and discrete heterogeneity from single particle cryo-EM. *J Struct Biol* **213**, 107702 (2021). https://doi.org/10.1016/j.jsb.2021.107702
24 Enemark, E. J. & Joshua-Tor, L. Mechanism of DNA translocation in a replicative hexameric helicase. *Nature* **442**, 270-275 (2006). https://doi.org/10.1038/nature04943
25 He, J., Li, T. & Huang, S. Y. Improvement of cryo-EM maps by simultaneous local and non-local deep learning. *Nat Commun* **14**, 3217 (2023). https://doi.org/10.1038/s41467-023-39031-1
26 Croll, T. I. ISOLDE: a physically realistic environment for model building into low-resolution electron-density maps. *Acta Crystallogr D Struct Biol* **74**, 519-530 (2018). https://doi.org/10.1107/S2059798318002425
27 Pintilie, G. *et al.* Measurement of atom resolvability in cryo-EM maps with Q-scores. *Nat Methods* **17**, 328-334 (2020). https://doi.org/10.1038/s41592-020-0731-1
28 Shein, M. *et al.* Characterizing ATP processing by the AAA+ protein p97 at the atomic level. *Nat Chem* (2024). https://doi.org/10.1038/s41557-024-01440-0
29 Enemark, E. J. & Joshua-Tor, L. On helicases and other motor proteins. *Curr Opin Struct Biol* **18**, 243-257 (2008). https://doi.org/10.1016/j.sbi.2008.01.007
30 Fernandez, A. J. & Berger, J. M. Mechanisms of hexameric helicases. *Crit Rev Biochem Mol Biol* **56**, 621-639 (2021). https://doi.org/10.1080/10409238.2021.1954597
31 O'Donnell, M. E. & Li, H. The ring-shaped hexameric helicases that function at DNA replication forks. *Nat Struct Mol Biol* **25**, 122-130 (2018). https://doi.org/10.1038/s41594-018-0024-x
32 Gao, Y. *et al.* Structures and operating principles of the replisome. *Science* **363** (2019). https://doi.org/10.1126/science.aav7003
33 Itsathitphaisarn, O., Wing, R. A., Eliason, W. K., Wang, J. & Steitz, T. A. The hexameric helicase DnaB adopts a nonplanar conformation during translocation. *Cell* **151**, 267-277 (2012). https://doi.org/10.1016/j.cell.2012.09.014

34 Meagher, M., Epling, L. B. & Enemark, E. J. DNA translocation mechanism of the MCM complex and implications for replication initiation. *Nat Commun* **10**, 3117 (2019). https://doi.org/10.1038/s41467-019-11074-3
35 Eickhoff, P. *et al.* Molecular Basis for ATP-Hydrolysis-Driven DNA Translocation by the CMG Helicase of the Eukaryotic Replisome. *Cell Rep* **28**, 2673-2688 e2678 (2019). https://doi.org/10.1016/j.celrep.2019.07.104
36 Yuan, Z. *et al.* Structure of the eukaryotic replicative CMG helicase suggests a pumpjack motion for translocation. *Nat Struct Mol Biol* **23**, 217-224 (2016). https://doi.org/10.1038/nsmb.3170
37 Pellegrini, L. The CMG DNA helicase and the core replisome. *Curr Opin Struct Biol* **81**, 102612 (2023). https://doi.org/10.1016/j.sbi.2023.102612
38 Xu, Z. *et al.* Synergism between CMG helicase and leading strand DNA polymerase at replication fork. *Nat Commun* **14**, 5849 (2023). https://doi.org/10.1038/s41467-023-41506-0
39 Rzechorzek, N. J., Hardwick, S. W., Jatikusumo, V. A., Chirgadze, D. Y. & Pellegrini, L. CryoEM structures of human CMG-ATPgammaS-DNA and CMG-AND-1 complexes. *Nucleic Acids Res* **48**, 6980-6995 (2020). https://doi.org/10.1093/nar/gkaa429
40 Borowiec, J. A. & Hurwitz, J. Localized melting and structural changes in the SV40 origin of replication induced by T-antigen. *EMBO J* **7**, 3149-3158 (1988). https://doi.org/10.1002/j.1460-2075.1988.tb03182.x
41 Wiekowski, M., Schwarz, M. W. & Stahl, H. Simian virus 40 large T antigen DNA helicase. Characterization of the ATPase-dependent DNA unwinding activity and its substrate requirements. *J Biol Chem* **263**, 436-442 (1988).
42 Remus, D. *et al.* Concerted loading of Mcm2-7 double hexamers around DNA during DNA replication origin licensing. *Cell* **139**, 719-730 (2009). https://doi.org/10.1016/j.cell.2009.10.015
43 Greiwe, J. F. *et al.* Structural mechanism for the selective phosphorylation of DNA-loaded MCM double hexamers by the Dbf4-dependent kinase. *Nat Struct Mol Biol* **29**, 10-20 (2022). https://doi.org/10.1038/s41594-021-00698-z
44 Lewis, J. S. *et al.* Mechanism of replication origin melting nucleated by CMG helicase assembly. *Nature* **606**, 1007-1014 (2022). https://doi.org/10.1038/s41586-022-04829-4
45 Yuan, Z. *et al.* DNA unwinding mechanism of a eukaryotic replicative CMG helicase. *Nat Commun* **11**, 688 (2020). https://doi.org/10.1038/s41467-020-14577-6
46 Mayle, R. *et al.* Mcm10 has potent strand-annealing activity and limits translocase-mediated fork regression. *Proc Natl Acad Sci U S A* **116**, 798-803 (2019). https://doi.org/10.1073/pnas.1819107116
47 Langston, L. D., Georgescu, R. E. & O'Donnell, M. E. Mechanism of eukaryotic origin unwinding is a dual helicase DNA shearing process. *Proc Natl Acad Sci U S A* **120**, e2316466120 (2023). https://doi.org/10.1073/pnas.2316466120
48 Li, J. *et al.* The human pre-replication complex is an open complex. *Cell* **186**, 98-111 e121 (2023). https://doi.org/10.1016/j.cell.2022.12.008
49 SenGupta, D. J. & Borowiec, J. A. Strand-specific recognition of a synthetic DNA replication fork by the SV40 large tumor antigen. *Science* **256**, 1656-1661 (1992). https://doi.org/10.1126/science.256.5064.1656

50 Bokori-Brown, M. *et al.* Cryo-EM structure of lysenin pore elucidates membrane insertion by an aerolysin family protein. *Nat Commun* **7**, 11293 (2016). https://doi.org/10.1038/ncomms11293
51 Kimanius, D., Dong, L., Sharov, G., Nakane, T. & Scheres, S. H. W. New tools for automated cryo-EM single-particle analysis in RELION-4.0. *Biochem J* **478**, 4169-4185 (2021). https://doi.org/10.1042/BCJ20210708
52 Zheng, S. Q. *et al.* MotionCor2: anisotropic correction of beam-induced motion for improved cryo-electron microscopy. *Nat Methods* **14**, 331-332 (2017). https://doi.org/10.1038/nmeth.4193
53 Rohou, A. & Grigorieff, N. CTFFIND4: Fast and accurate defocus estimation from electron micrographs. *J Struct Biol* **192**, 216-221 (2015). https://doi.org/10.1016/j.jsb.2015.08.008
54 Bepler, T. *et al.* Positive-unlabeled convolutional neural networks for particle picking in cryo-electron micrographs. *Nat Methods* **16**, 1153-1160 (2019). https://doi.org/10.1038/s41592-019-0575-8
55 Tan, Y. Z. *et al.* Addressing preferred specimen orientation in single-particle cryo-EM through tilting. *Nat Methods* **14**, 793-796 (2017). https://doi.org/10.1038/nmeth.4347
56 Jakobi, A. J., Wilmanns, M. & Sachse, C. Model-based local density sharpening of cryo-EM maps. *Elife* **6** (2017). https://doi.org/10.7554/eLife.27131
57 Jumper, J. *et al.* Highly accurate protein structure prediction with AlphaFold. *Nature* **596**, 583-589 (2021). https://doi.org/10.1038/s41586-021-03819-2
58 Meng, E. C., Pettersen, E. F., Couch, G. S., Huang, C. C. & Ferrin, T. E. Tools for integrated sequence-structure analysis with UCSF Chimera. *BMC Bioinformatics* **7**, 339 (2006). https://doi.org/10.1186/1471-2105-7-339
59 Kabsch, W. & Sander, C. Dictionary of protein secondary structure: pattern recognition of hydrogen-bonded and geometrical features. *Biopolymers* **22**, 2577-2637 (1983). https://doi.org/10.1002/bip.360221211
60 Pettersen, E. F. *et al.* UCSF Chimera--a visualization system for exploratory research and analysis. *J Comput Chem* **25**, 1605-1612 (2004). https://doi.org/10.1002/jcc.20084
61 Meng, E. C. *et al.* UCSF ChimeraX: Tools for structure building and analysis. *Protein Sci* **32**, e4792 (2023). https://doi.org/10.1002/pro.4792

Acknowledgements

This research was supported by King Abdullah University of Science and Technology (KAUST) through core funding (to S.M.H. and A.D.B.) and the Competitive Research Award Grant CRG8 URF/1/4036-01-01 (to S.M.H. and A.D.B.). We thank past and present members of the S.M.H. and A.D.B. laboratories for help and discussions; Lingyun Zhao, Ashraf Al-Alamoudi, Alessandro Genovese and Rachid Sougrat for their assistance

at the Imaging and Characterization Core Lab at KAUST. We acknowledge The Midlands Regional Cryo-EM Facility at the Leicester Institute of Structural and Chemical Biology (LISCB), major funding from MRC (MC_PC_17136). We thank T.J. Ragan (LISCB, University of Leicester) for his advise on data processing.

Supplementary Information

SI Guide

Internal unwinding mechanism of a replicative helicase

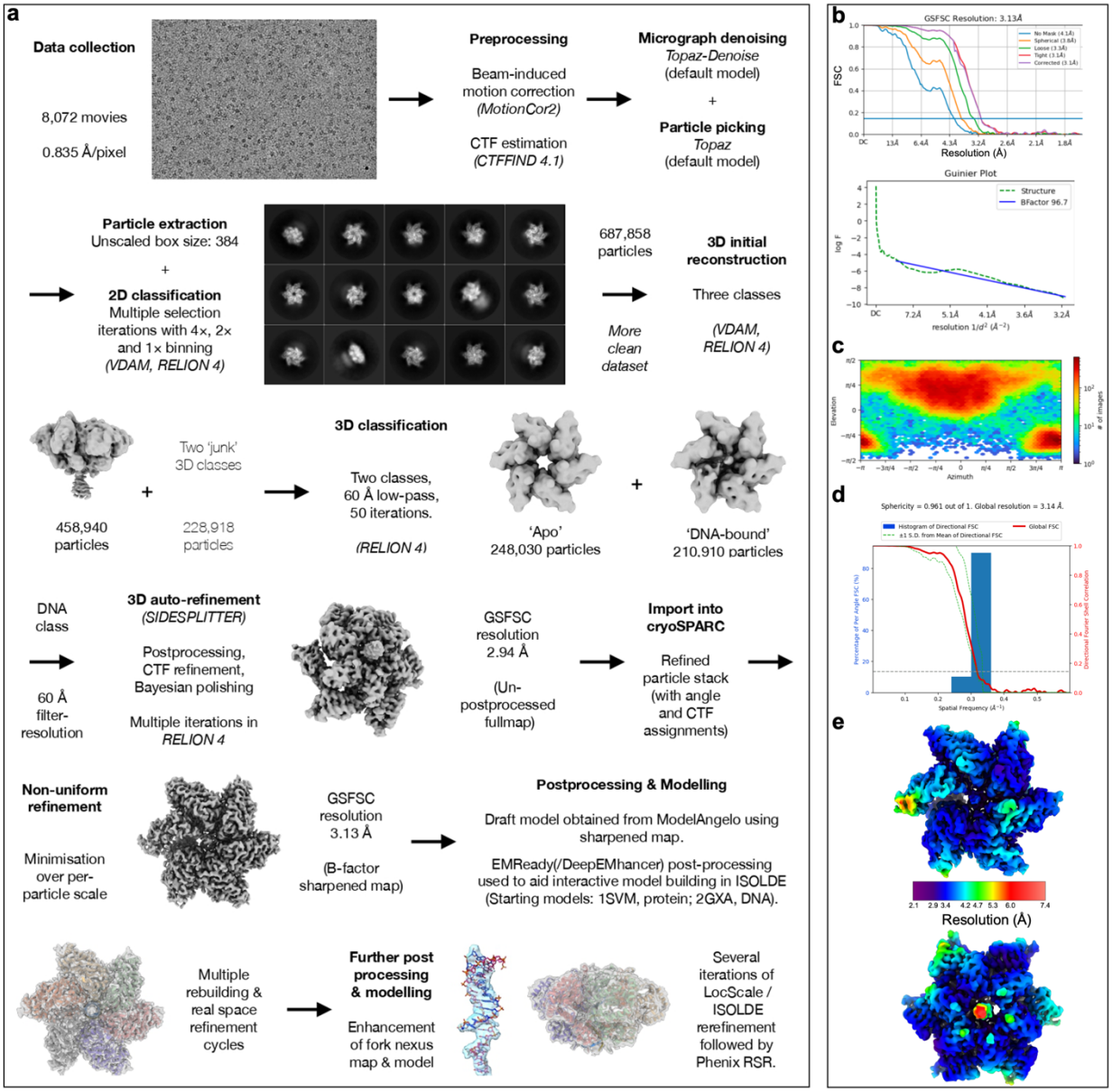
Taha Shahid[1,2], Muhammad Tehseen[1], Lubna Alhudhali[1], Alice Clark[1], Christos G. Savva[1], Samir M. Hamdan[1] and Alfredo De Biasio[1,2]

[1]Bioscience Program, Division of Biological and Environmental Sciences and Engineering, King Abdullah University of Science and Technology, Thuwal 23955, Saudi Arabia.

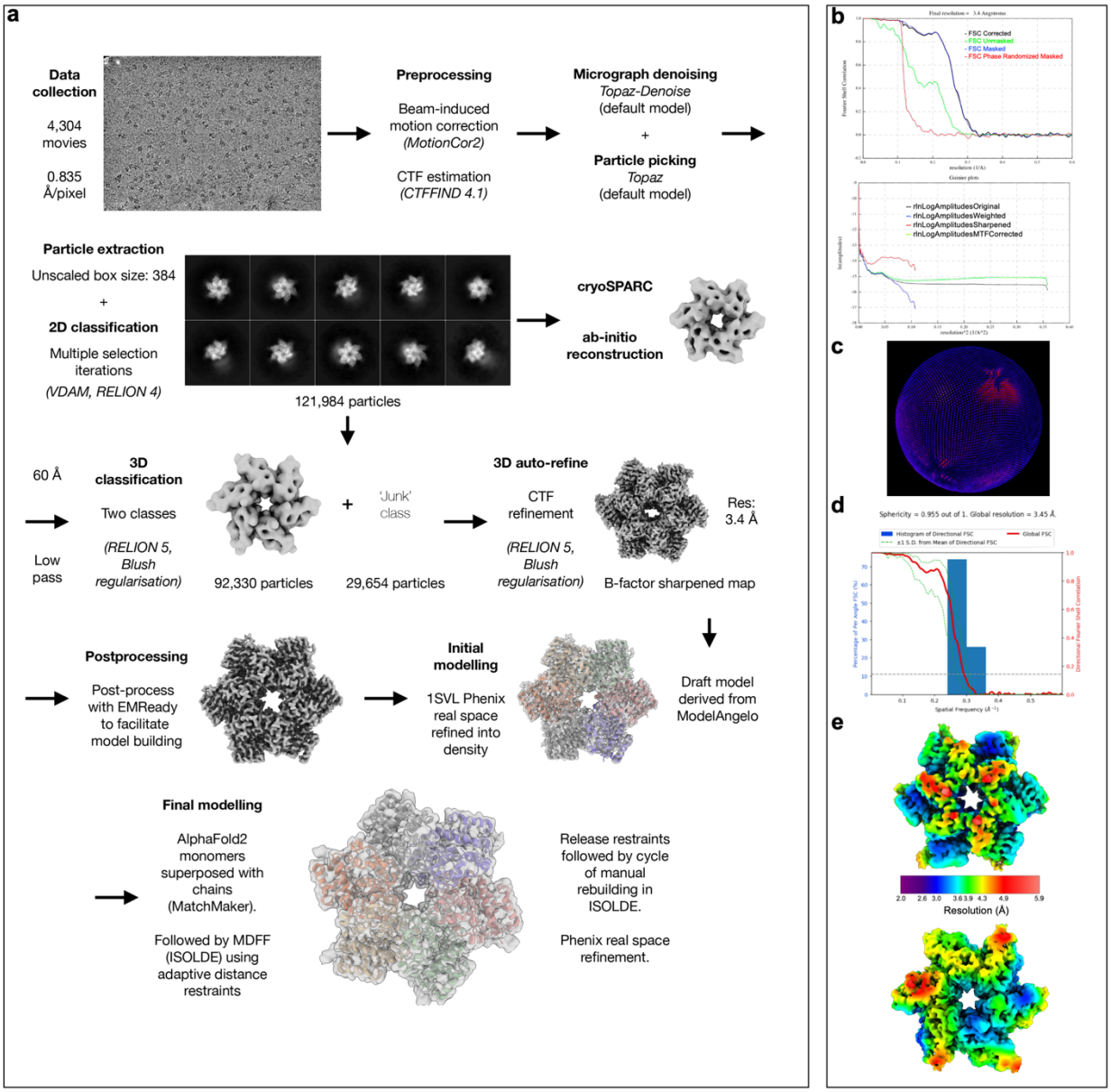
[2]Leicester Institute of Structural & Chemical Biology and Department of Molecular & Cell Biology, University of Leicester, Lancaster Rd, Leicester LE1 7HB, UK.

Table of contents

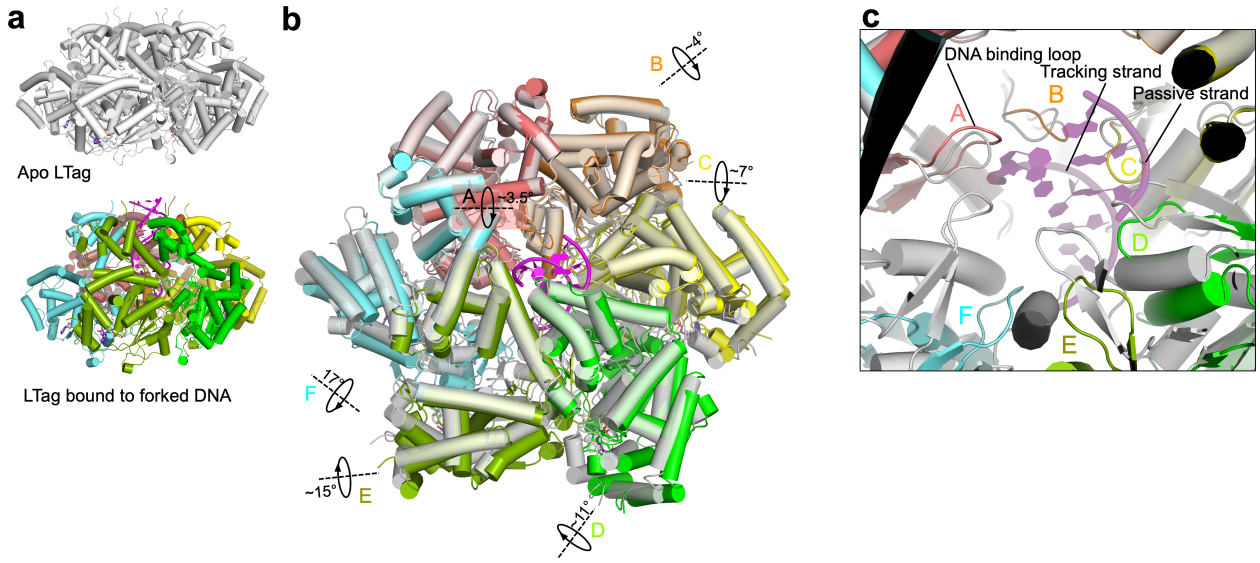
Extended Data Figure 1
Extended Data Figure 2
Extended Data Figure 3
Extended Data Figure 4
Extended Data Figure 5
Extended Data Figure 6
Extended Data Figure 7
Extended Data Figure 8
Extended Data Figure 9
Extended Data Figure 10
Extended Data Table 1
Extended Data Table 2
Supplementary Video 1
Supplementary Video 2
Supplementary Video 3
Supplementary Video 4
Supplementary Video 5
Supplementary Video 6
Supplementary Video 7
Supplementary Video 8
Extended Data References



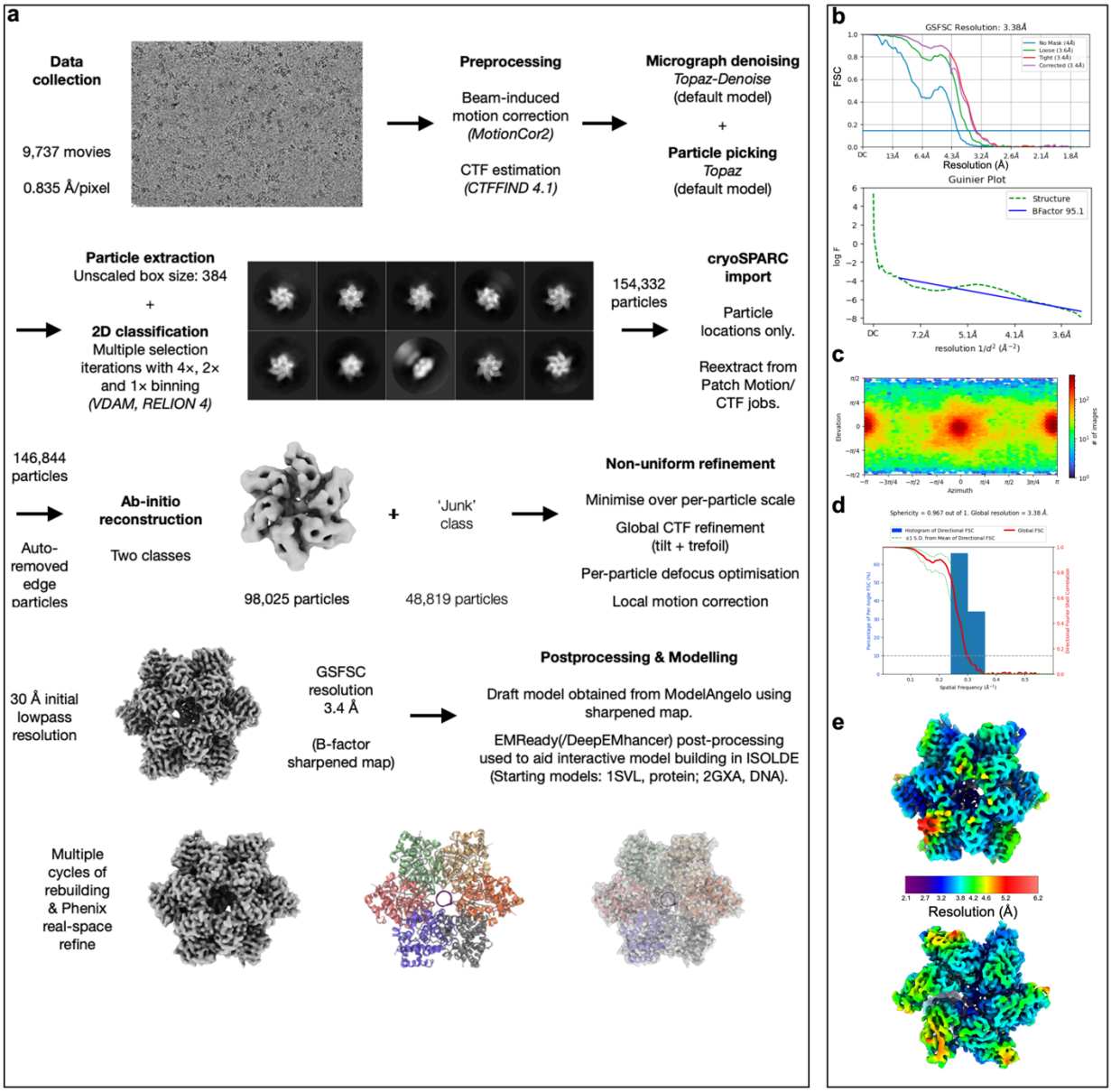
Extended Data Figure 1. *Cryo-EM of the LTag–ATP–DNA complex.* **a,** Workflow schematics for cryo-EM data collection, processing, and modelling. Flows continue from left hand side of following row unless otherwise indicated **b,** Gold-standard Fourier shell correlation and Gunier plots **c,** Angular distribution of projections. Resolution is estimated using the 0.143 criterion. **d,** Map anisotropy analysis computed by 3DFSC[1]. **e,** Two views of the cryo-EM map colored by local resolution.



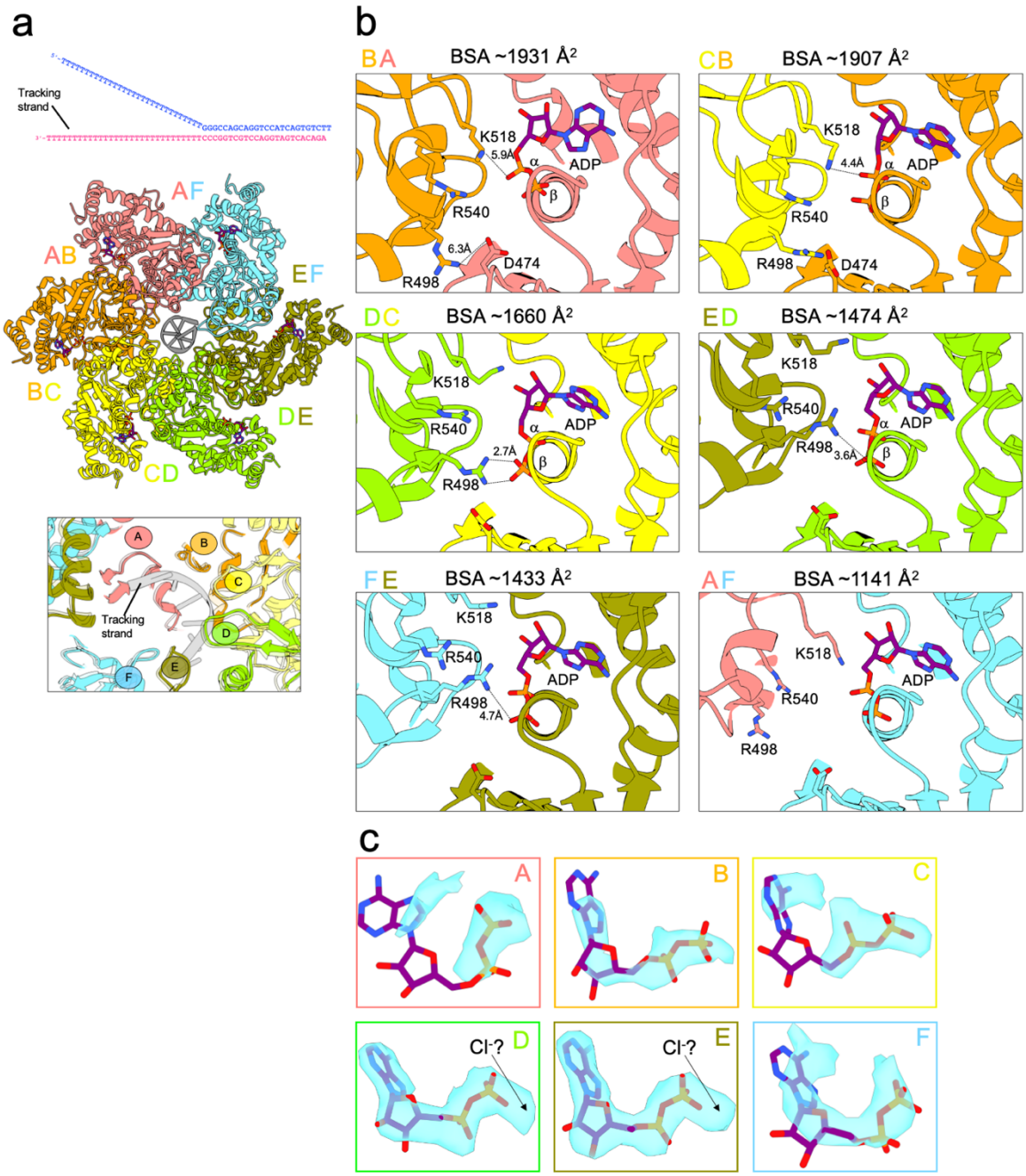
Extended Data Figure 2. *Cryo-EM of the LTag–ATP (APO) complex.* **a,** Workflow schematics for cryo-EM data collection, processing, and modelling. Flows continue from left hand side of following row unless otherwise indicated **b,** Gold-standard Fourier shell correlation and Gunier plots **c,** Angular distribution of projections. Resolution is estimated using the 0.143 criterion. **d,** Map anisotropy analysis computed by 3DFSC[1]. **e,** Two views of the cryo-EM map colored by local resolution.



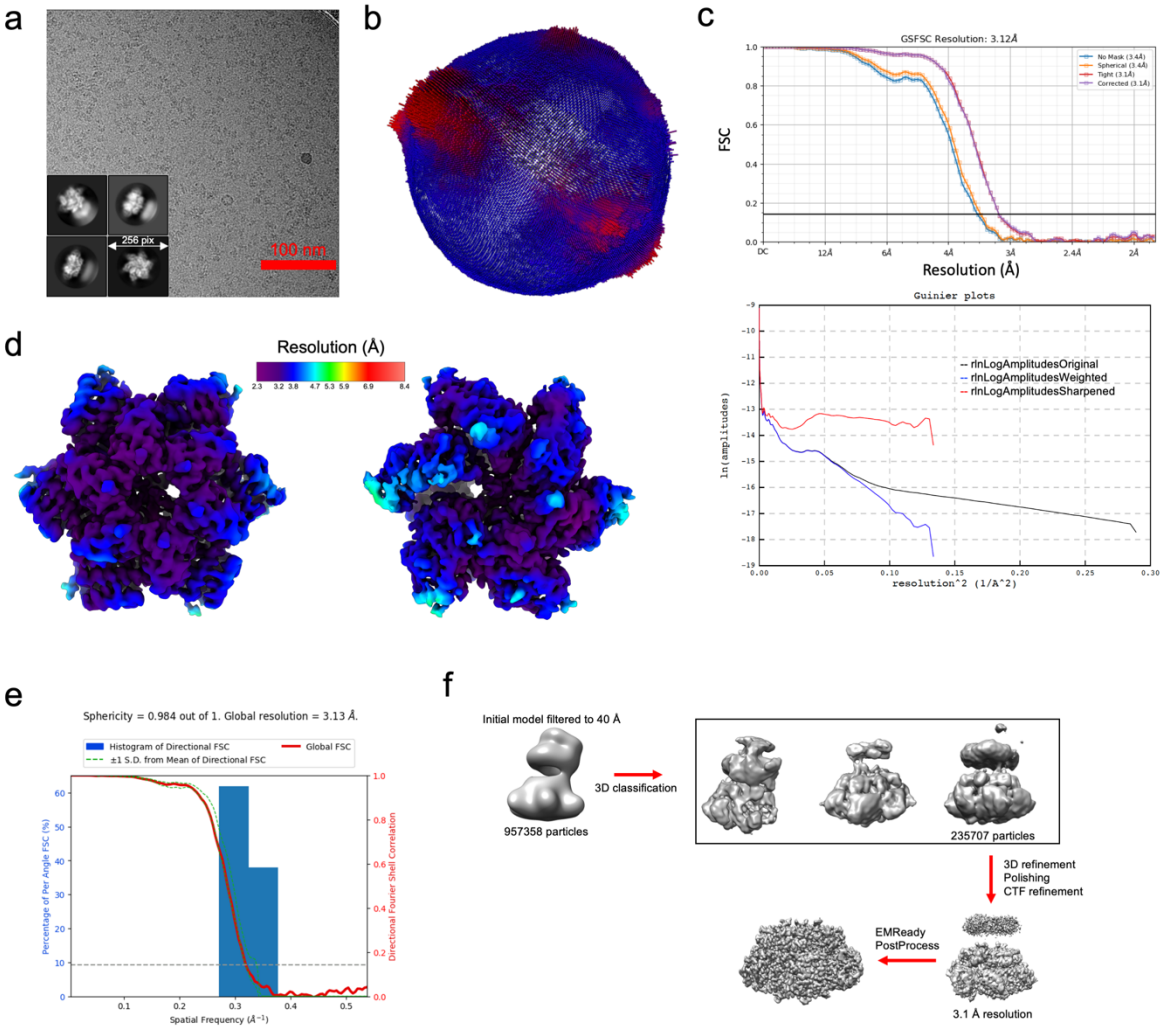
Extended Data Figure 3. *Structural comparison of apo LTag and LTag bound to forked DNA.* **a,** Comparative cryo-EM models of the apo (unbound) LTag and LTag bound to forked DNA, **b,** Superposition of the models from **(a)**, illustrating the incremental rotations of the LTag C-tier domains upon DNA binding, attributed to the engagement of DNA-binding loops with the tracking strand. **c,** Close-up view of the LTag central pore in the superimposed structures from **(b)**, comparing the DNA-binding loops. In the apo LTag, loops are shown in an equiplanar arrangement, contrasting with a staircase configuration in the DNA-bound LTag, indicating significant structural reorganization upon DNA engagement.



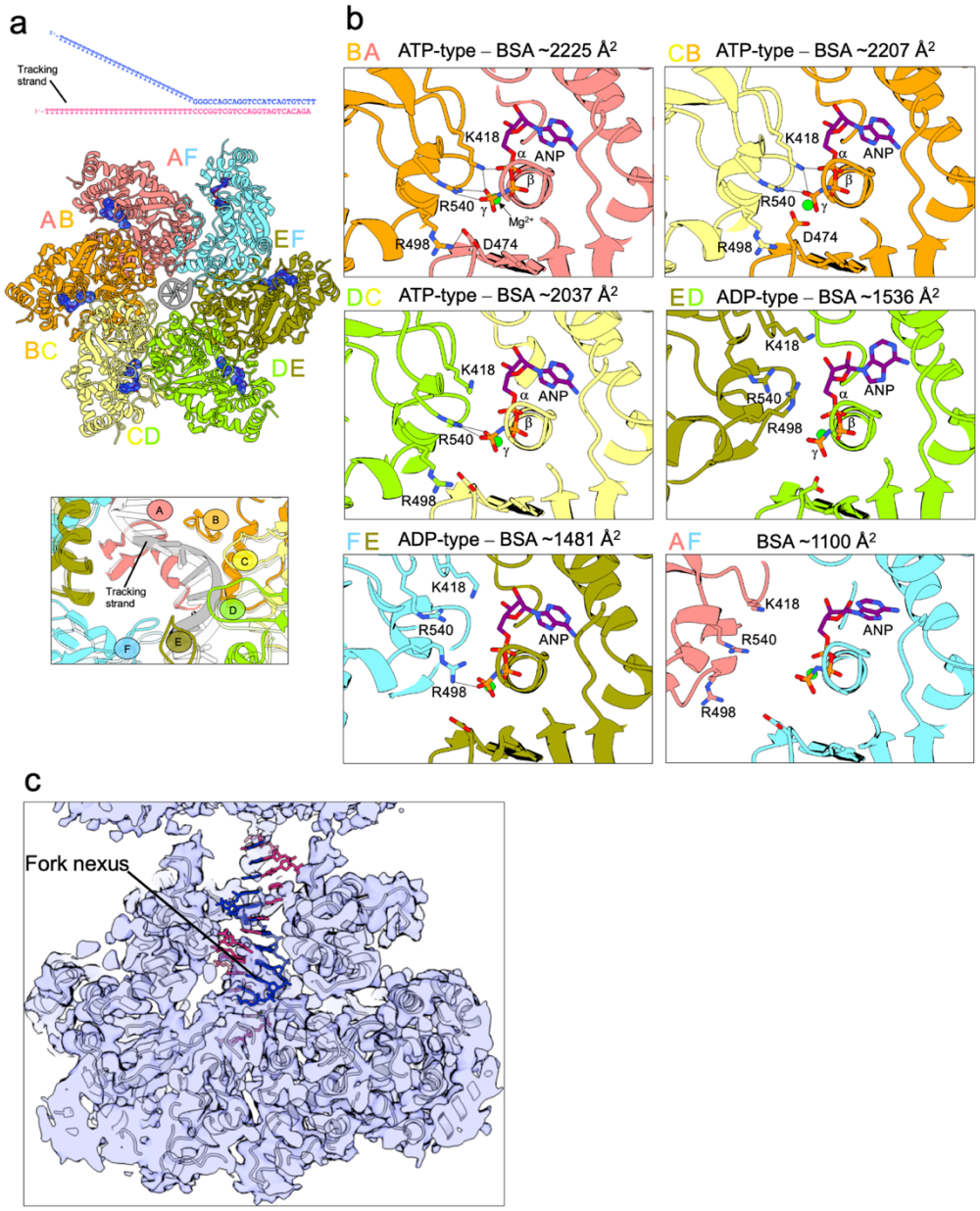
Extended Data Figure 4. *Cryo-EM of the LTag–ADP–DNA complex.* **a,** Workflow schematics for cryo-EM data collection, processing, and modelling. Flows continue from left hand side of following row unless otherwise indicated. **b,** Gold-standard Fourier shell correlation and Gunier plots **c,** Angular distribution of projections. Resolution is estimated using the 0.143 criterion. **d,** Map anisotropy analysis computed by 3DFSC[1]. **e,** Two views of the cryo-EM map colored by local resolution.



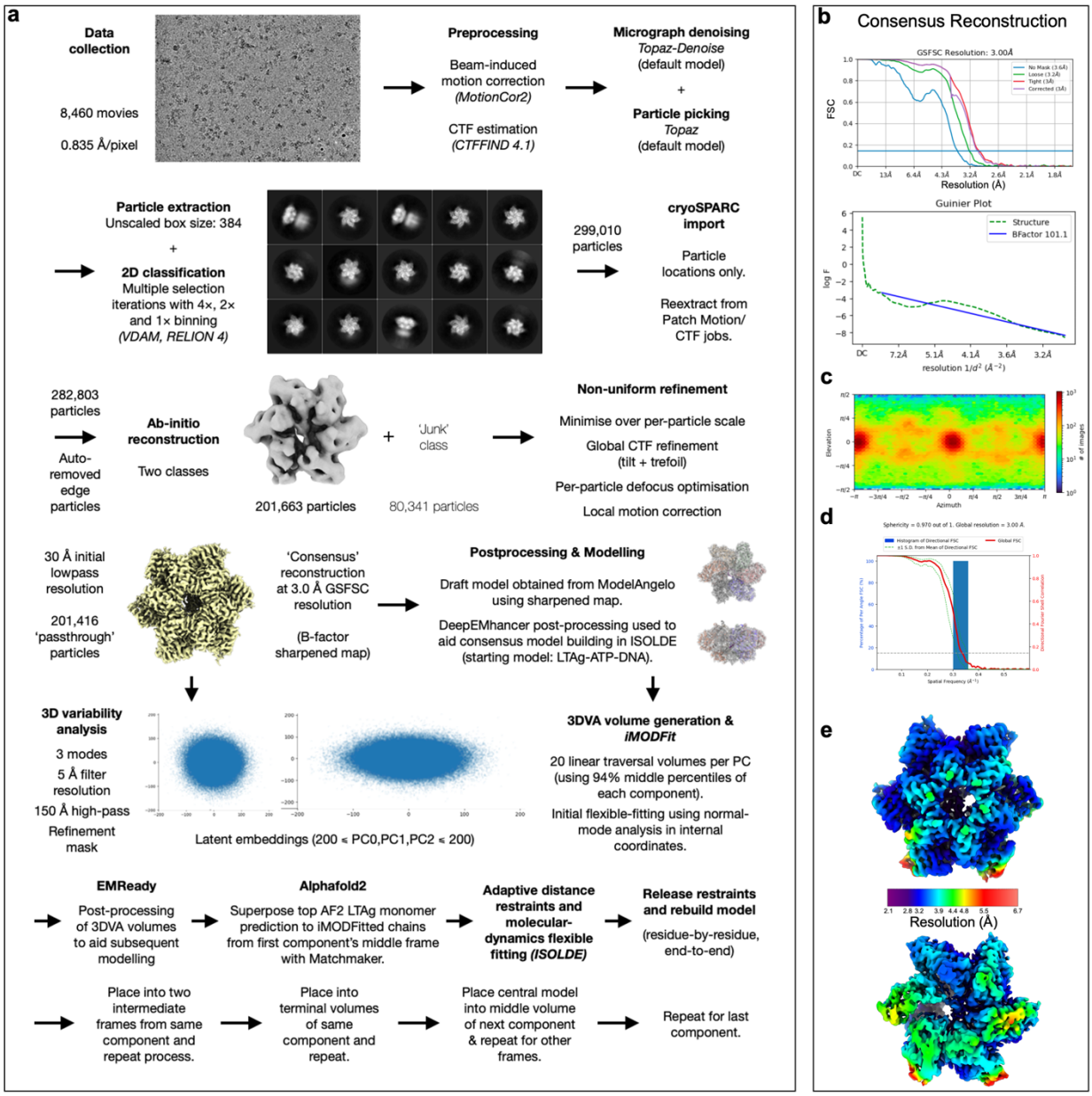
Extended Data Figure 5. *Inter-subunit interactions and nucleotide binding in the LTag–ADP–forked DNA complex.* **a,** Bottom view of the cryo-EM model. The sequence of the DNA substrate is shown above. The lower inset shows details of the DNA binding loop staircase and DNA. **b,** Inter-subunit interactions at the nucleotide binding pocket. Calculated Buried Surface Areas (BSA) for each interface are shown, along with the assigned interface type. Hydrogen bonds are shown as black dotted lines. **c,** Map density around the ATP molecules in the various LTag subunits. Density at γ-phosphate position in subunits D and E may be attributed to the presence of a chloride atom.



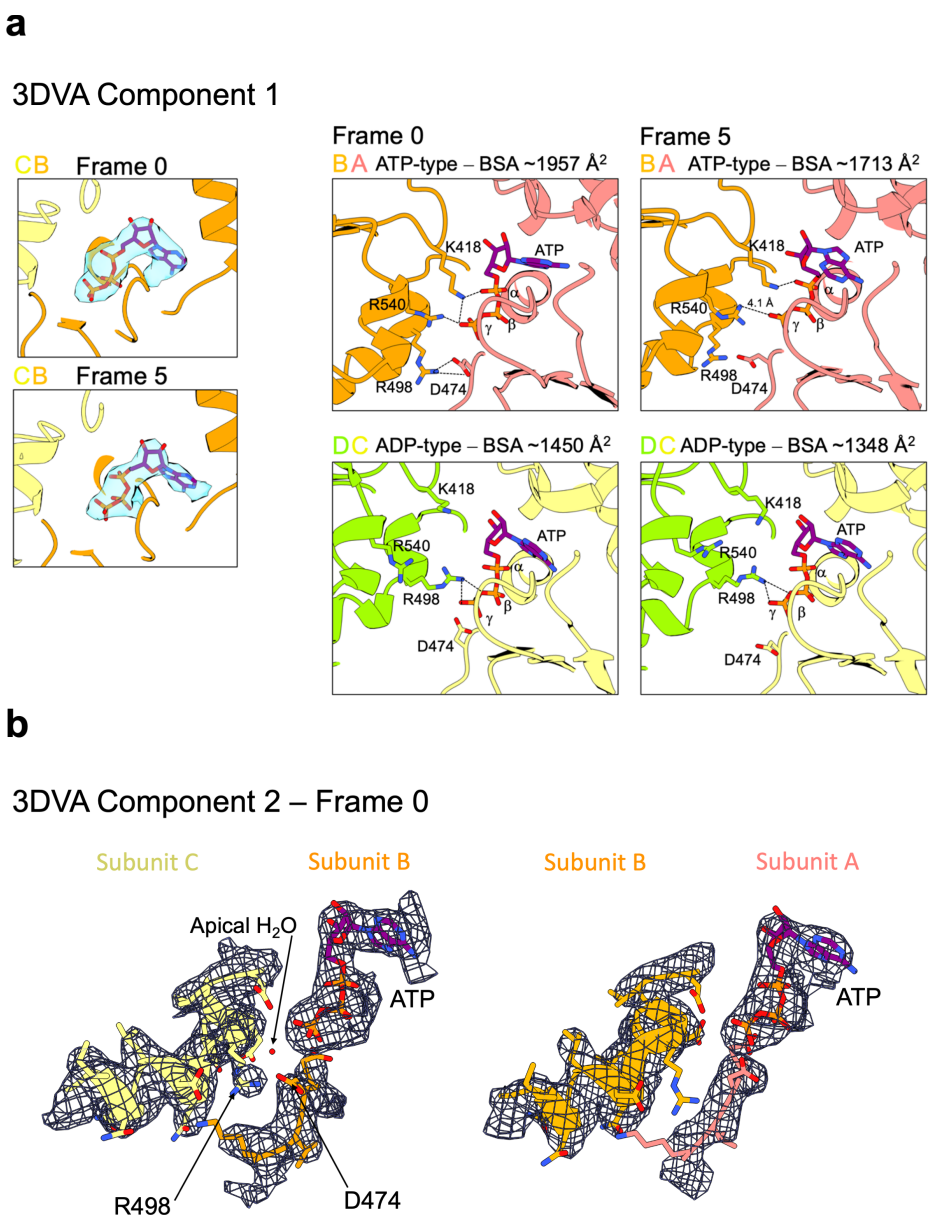
Extended Data Figure 6. *Cryo-EM of the LTag–ANPPNP/Mg^{2+}–DNA complex.* **a,** Electron micrograph (aligned sum) acquired using a Falcon 4 direct electron detector and representative 2D class averages. **b,** Angular distribution of projections. **c,** Gold-standard Fourier shell correlation and Gunier plots. Resolution is estimated using the 0.143 criterion. **d,** Two views of the cryo-EM map colored by local resolution. **e,** Map anisotropy analysis computed by 3DFSC[1]. **d,** Overview of image processing.



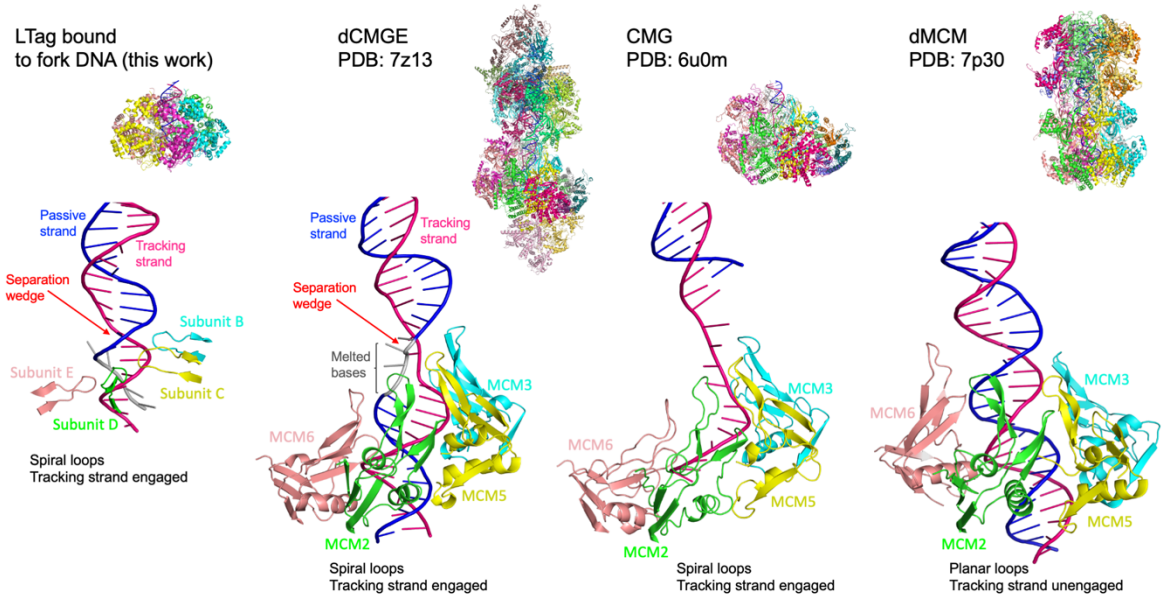
Extended Data Figure 7. *Inter-subunit interactions and nucleotide binding in the LTag–AMP/PNP–forked DNA complex.* **a,** Bottom view of the cryo-EM model. The sequence of the DNA substrate is shown above. The lower inset shows details of the DNA binding loop staircase and DNA. The DNA shown in transparency corresponds to the position in the LTag structure reconstituted with ATP. **b,** Analysis of the nucleotide-binding pocket shows inter-subunit interactions with calculated Buried Surface Areas (BSA) and hydrogen bonds (black dotted lines). **c,** A cut-through of the cryo-EM map displays density at the DNA nexus, with the DNA model (in stick form) aligning with the structure of LTag bound to forked DNA and ATP.



Extended Data Figure 8. *Cryo-EM of the LTag–ATP/Mg^{2+}–DNA complex.* **a,** Workflow schematics for cryo-EM data collection, processing, and modelling. Flows continue from left hand side of following row unless otherwise indicated. **b,** Gold-standard Fourier shell correlation and Gunier plots for the consensus reconstruction **c,** Angular distribution of projections. Resolution is estimated using the 0.143 criterion. **d,** Map anisotropy analysis computed by 3DFSC[1]. **e,** Two views of the cryo-EM map colored by local resolution.



Extended Data Figure 9. *Analysis of the nucleotide binding pocket in the first and second 3DVA component trajectories of LTag translocating on forked DNA*. **a,** *Left:* Map density around the ATP molecule at the B/C interface in the initial and final frames of the first 3DVA component. Density at γ-phosphate position persists throughout the trajectory despite a full breakage of the interface, suggesting that γ-phosphate may be retained at the pocket post-hydrolysis. *Right:* Interactions at the nucleotide binding pocket between the A and B or C and D subunits in the initial and final frames of the first 3DVA component. Calculated Buried Surface Areas (BSA) are shown, along with the assigned interface type. Hydrogen bonds are shown as black dotted lines. The analysis reveals a moderate disruption of the A/B interface and no disruption of the C/D interface. **b,** Maps and models of the BC and AB interfaces at the initial frame of second 3DVA component, showing order of the R498 side chain in C but not B subunits. The position of the apical water molecule, inferred from LTag-ATP crystal structure (PDB ID 1SVM)[2], is indicated.

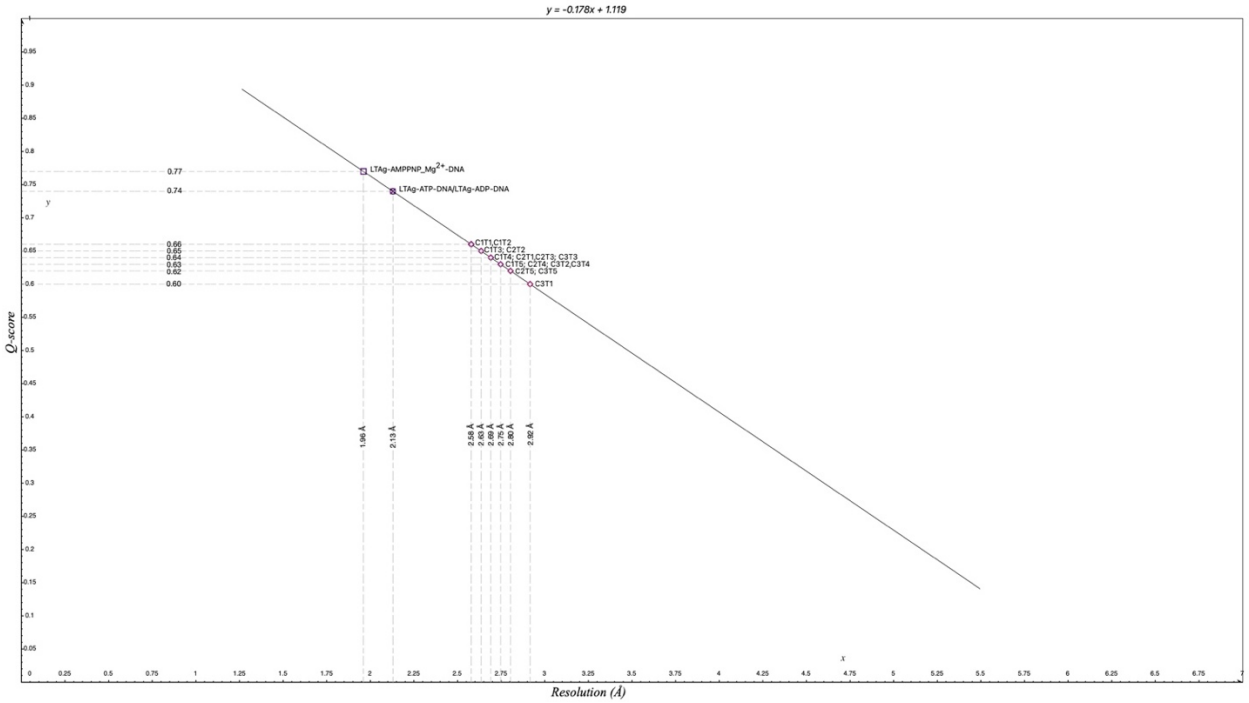


Extended Data Figure 10. *Comparison of internal melting mechanisms in LTag and CMG Helicases*. This figure highlights the internal DNA separation mechanism within the LTag helicase bound to forked DNA and its resemblance to the yeast CMG double hexamer structure during DNA melting at the origin[3] (dCMGE; PDB 7z13). In both cases, the formation of a separation wedge is facilitated by ATP binding, the staircase arrangement of DNA binding loops, and the engagement of these loops with the tracking strand. These elements collectively drive strand separation within the helicase inner chamber, positioning the helicase for ATP-powered translocation. This configuration is mirrored in the actively translocating CMG structure[4] (PDB 6u0m), showcasing the conserved mechanism of action. Conversely, the pre-CMG formation MCM double hexamer[5] (dMCM; PDB 7p30) displays planar DNA binding loops and a disengaged tracking strand, indicative of a state not yet primed for ATP-dependent movement.

	LTAg-ATP	LTAg-ATP-DNA (LocScale)	LTAg-ADP-DNA	Consensus LTAg-ATP/Mg^{2+}-DNA (deepEMhancer highRes)	LTAg-AMPPNP/ Mg^{2+}-DNA
Data collection and processing					
Magnification					
Voltage (kV)	300	300	300	300	300
Electron exposure (e–/Å^2)	50	50	50	50	39
Defocus range (μm)	-1.0 – 2.5	-1.0 – 2.5	-1.0 – 2.5	-1.0 – 2.5	-1.5 – -2.5
Pixel size (Å)	0.835	0.835	0.835	0.835	0.93
Symmetry imposed	C1	C1	C1	C1	C1
Initial particle images (no.)					
Final particle images (no.)	92,330	210,910	95,587	201,416	235,707
Map resolution (Å)	3.4	3.1	3.4	3.0	3.2
FSC threshold	0.143	0.143	0.143	0.143	0.143
Map resolution range (Å)	2.1 – 5.9	2.1 – 7.4	2.1 – 6.2	2.1 – 6.7	2.3 – 8.4
Refinement					
Initial model used (PDB code)	LTAg-ATP-DNA	1SVM	1SVL	LTAg-ATP-DNA	LTAg-ATP-DNA
Model resolution (Å)	2.1	3.3	2.0	3.0	1.5
FSC threshold	0.5	0.5	0.5	0.5	0.5
Model resolution range (Å)	1.3 – 2	2.9 – 3.3	1.3 – 2	1.3 – 3.0	1.0 – 1.5
Map sharpening *B* factor (Å^2)	-92.9	-96.7	-95.1	-101.1	-127.9
Model composition					
Non-hydrogen atoms	17,778	18,349	17,894	17,864	17,924
Protein residues	2172	2172	2172	2166	2172
Ligands	6	6	6	6	6
B factors (Å^2)					
Protein	75.14	127.29	74.11	73.73	59.53
Ligand	95.70	154.74	101.34	84.06	78.98
R.m.s. deviations					
Bond lengths (Å)	0.006	0.005	0.006	0.007	0.006
Bond angles (°)	1.222	1.152	1.031	1.224	0.960
Validation					
MolProbity score	0.79	0.78	0.78	0.83	0.76
Clashscore	0.50	0.46	0.36	0.56	0.86
Poor rotamers (%)	0	0	0	0	0
Ramachandran plot					
Favored (%)	97.41	97.41	97.18	97.26	98.01
Allowed (%)	2.59	2.59	2.82	2.74	1.99
Disallowed (%)	0	0	0	0	0

*EMReady postprocessed maps

Extended Data Table 1 Cryo-EM data collection, refinement and validation statistics



Extended Data Table 2. Q-scores[6] of modelled structures against respective maps postprocessed by EMReady[7] (horizontal dotted lines). C1, C2, C3 refer to 3DVA principal components. T1, T2, T3, T4, T5 refer to modelled volumes from each component, in order of reaction sequence. Vertical dotted lines indicate GSFSC resolutions ordinarily corresponding to said Q-scores, underscoring quality of modelling. Solid line represents linear fit of average Q-score versus reported resolution of deposited map-model pairs in EMDB/PDB[6].

Supplementary Video 1: Morphing between five 3DVA volumes across the first component trajectory. Volumes are colored according to the different LTag subunits or DNA.

Supplementary Video 2: Morphing between five 3DVA volumes across the second component trajectory. Volumes are colored according to the specific LTag subunits and DNA.

Supplementary Video 3: Morphing between five 3DVA volumes across the third component trajectory. Volumes are colored according to the specific LTag subunits and DNA.

Supplementary Video 4: Morphing between five 3DVA molecular models across the first component trajectory. This close-up highlights the net translocation of DNA across the LTag channel and the movement of the DNA binding loops. LTag subunits are shown as ribbons, DNA and the side chain of H513 as sticks.

Supplementary Video 5: Molecular dynamics movie demonstrating a metastable hydrogen phosphate + ADP in post-hydrolysis configuration inside an “ADP-type" nucleotide pocket. The hydrolyzed phosphate is stabilised by surrounding positive/polar residues. The nucleotide pocket and relative cryo-EM map corresponds to the C subunit in the last frame of the first 3DVA component. The simulation was carried out with ISOLDE [8].

Supplementary Video 6: Morphing between five 3DVA molecular models across the second component trajectory. This close-up highlights the net translocation of DNA across the LTag channel and the upward movement of the DNA binding loop of the F subunit (colored in light blue), signifying a preparatory step for translocation restart. LTag subunits are shown as ribbons, DNA and the side chain of H513 as sticks.

Supplementary Video 7: Morphing between five 3DVA molecular models across the third component trajectory. This close-up highlights that there is no net translocation of DNA (see text for details). LTag subunits are shown as ribbons, DNA and the side chain of H513 as sticks.

Supplementary Video 8: The animation illustrates the predicted process of DNA unwinding by LTag throughout a full translocation cycle, generated by morphing three structures. The first and second structures represent the initial and final frames of the second 3DVA component, respectively. The third structure is based on the initial frame, but with the LTag subunits rotated one step around the six-fold symmetry axis of the N-tier. The DNA duplex segment is modeled after the static structure of LTag bound to forked DNA and ATP, but without Mg^{2+}. The translocation and unwinding process is initiated by ATP hydrolysis at the hexamer most constrained interface (between subunits B and C, colored orange and yellow), serving as an 'entropy switch'. This hydrolysis triggers coordinated rotations of the C-tier domains in the subunits and adjustments in the DNA binding loops. These changes drive the DNA through the LTag internal chamber (with the tracking and passive strands colored magenta and blue, respectively). Simultaneously, the ADP-ATP exchange by the F subunit (colored in blue) raises the DNA binding loop of this subunit to the top of the staircase, forming the ATP interface with the A subunit. This re-establishes the staircase configuration, resulting in the melting of one base pair at the DNA nexus (shown in black). Notably, this model of internal unwinding features the tracking strand advancing through the DNA loops without rotation, while rotation occurs in the double-stranded DNA segment, as the next nucleobase at the top of the staircase is incorporated into the compacted helix of the tracking strand.

Extended Data References

1 Tan, Y. Z. *et al.* Addressing preferred specimen orientation in single-particle cryo-EM through tilting. *Nat Methods* **14**, 793-796 (2017). https://doi.org/10.1038/nmeth.4347

2 Gai, D., Zhao, R., Li, D., Finkielstein, C. V. & Chen, X. S. Mechanisms of conformational change for a replicative hexameric helicase of SV40 large tumor antigen. *Cell* **119**, 47-60 (2004). https://doi.org/10.1016/j.cell.2004.09.017

3 Lewis, J. S. *et al.* Mechanism of replication origin melting nucleated by CMG helicase assembly. *Nature* **606**, 1007-1014 (2022). https://doi.org/10.1038/s41586-022-04829-4

4 Yuan, Z. *et al.* DNA unwinding mechanism of a eukaryotic replicative CMG helicase. *Nat Commun* **11**, 688 (2020). https://doi.org/10.1038/s41467-020-14577-6

5 Greiwe, J. F. *et al.* Structural mechanism for the selective phosphorylation of DNA-loaded MCM double hexamers by the Dbf4-dependent kinase. *Nat Struct Mol Biol* **29**, 10-20 (2022). https://doi.org/10.1038/s41594-021-00698-z

6 Pintilie, G. *et al.* Measurement of atom resolvability in cryo-EM maps with Q-scores. *Nat Methods* **17**, 328-334 (2020). https://doi.org/10.1038/s41592-020-0731-1

7 He, J., Li, T. & Huang, S. Y. Improvement of cryo-EM maps by simultaneous local and non-local deep learning. *Nat Commun* **14**, 3217 (2023). https://doi.org/10.1038/s41467-023-39031-1

8 Croll, T. I. ISOLDE: a physically realistic environment for model building into low-resolution electron-density maps. *Acta Crystallogr D Struct Biol* **74**, 519-530 (2018). https://doi.org/10.1107/S2059798318002425