Rep structures can be tuned by ionicity via metastable intermediates in the absence of DNA

Jamieson A L Howard[1], Benjamin Ambrose[2,3], Mahmoud A S Abdelhamid[2], Antoinette Alevropoulos-Borrill[1], Ayesha Ejaz[4], Lara Dresser[1], Maria Dienerowitz[5], Steven D Quinn[1,6], Allison H Squires[7,8], Agnes Noy[1,6], Timothy D Craggs[2], and Mark C Leake[1,6,9] †

[1] School of Physics, Engineering and Technology, University of York, York, YO10 5DD, UK

[2] Department of Chemistry, University of Sheffield, Sheffield S3 7HF, U.K.

[3] Current address: Single Molecule Imaging Group, MRC-London Institute of Medical Sciences, London, W12 0HS, UK.

[4] Department of Chemistry, University of Chicago, Chicago, IL, USA.

[5] SciTec Department, Ernst-Abbe-Hochschule, University of Applied Sciences, Jena, Germany.

[6] York Biomedical Research Institute, University of York, York, YO10 5DD, UK.

[7] Pritzker School of Molecular Engineering, University of Chicago, Chicago, IL, USA.

[8] Institute for Biophysical Dynamics, University of Chicago, Chicago, IL, USA.

[9] Department of Biology, University of York, York, YO10 5DD, UK.

† For correspondence. Email mark.leake@york.ac.uk

Abstract

Many enzymes undergo conformational changes when bound to a substrate, however, structural plasticity in substrate-free enzymes, and potential associated benefits, is poorly understood. Here, we study the structural dynamics of Rep, an accessory helicase which undergoes conformational transitions in performing crucial roles in bacterial DNA replication, repair and recombination. Using time-correlated single-photon counting, fluorescence correlation spectroscopy, rapid single-molecule Förster resonance energy transfer, Anti-Brownian ELectrokinetic trapping and molecular dynamics simulations, we find that free Rep adopts four conformations, revealing that DNA absence confers structural plasticity beyond simple two-state binding. Our fast measurement technologies reveal hitherto hidden transient intermediate transitions between open and closed conformations, several of which are strongly dependent on sodium chloride concentration. Our findings support an alternative binding model for accessory helicases that utilises conformational plasticity to explore a multiplicity of structures whose landscape can be "tuned" by the ionic environment prior to locking-in upon DNA binding.

Keywords Helicases, single-molecule, FRET, structural dynamics, DNA replication, DNA-protein interactions, 1A superfamily

Introduction

Many enzymes are known to undergo structural conformational changes when bound to a substrate as part of their functional role during catalytic activity, however, the phenomenon of structural plasticity in the substrate-free enzyme, and its potential functional benefits, is poorly understood. Many DNA-binding proteins have been shown to exhibit such structural heterogeneity, one well studied class of enzymes being DNA polymerases[1–4]. Another class of DNA-binding proteins with

similar structural plasticity is helicases. These enzymes play crucial roles in DNA repair and the regulation and maintenance of genetic stability, with defects in viral/bacterial and human helicases linking to infectious diseases, genetic disorders and cancers[5]. Despite helicases being attractive targets for therapeutic intervention[6–8], key questions regarding how their structure is dynamically regulated remain. Although there are extensive data from biochemical, genetic, and structural tools new theoretical and experimental approaches are required to elucidate their time-resolved mechanistic behaviour. Helicases are ubiquitous molecular motors of nucleic acid metabolism, which utilize chemical potential energy derived from nucleoside triphosphate (NTP) hydrolysis (usually ATP) to translocate along a nucleic acid strand in a unidirectional manner[5,9–11] normally by 1-2 base pairs (bp) per NTP hydrolysed[5,9], and disrupt hydrogen bonds between complementary bp. Remodelling nucleic acids is not as simple as unwinding the double strands, however, as nucleic acids *in vivo* are bound by proteins that must simultaneously be removed at the time of strand separation[10]. It has generally been assumed that protein displacement occurs through coupling free energy produced during rounds of NTP hydrolysis and bp disruption, enabling helicases to generate force to push proteins along DNA[11–14]. However, the actual method of protein displacement remains to be elucidated.

This coating of nucleic acids with protein in live cells provides a necessity for accessory replicative helicases that aid the DNA replication machinery in progressing through nucleoprotein complexes which would otherwise act as a replication block[15–20]. Helicases are divided into six superfamilies based on shared sequence motifs[5], with *Escherichia coli* accessory replicative helicase Rep being an ATP-dependent superfamily 1A helicase that is essential under rapid growth conditions[16,17,21], structurally comparable to *E. coli* UvrD and *Geobacillus stearothermophilus* PcrA. Rep contains four subdomains (1A, 2A, 1B and 2B), in which 2B can undergo a substantial rotation about a hinge connected to 2A[5,22–28]. As with all superfamily 1A helicases, Rep has 3'-5' directionality and as such likely operates on the leading strand template displacing proteins ahead of the replication fork[16,17]. This activity is aided through a physical interaction via the C-terminus with the 5'-3' replicative helicase DnaB, which targets Rep to the replication fork[16,29–32] whilst also acting ahead of the bacterial replisome.

Although wild type Rep monomers are highly processive on single stranded DNA (ssDNA), monomeric unwinding of duplex DNA has remained undetected[23,33,34]. Monomeric unwinding of double-stranded DNA (dsDNA) is activated in the RepΔ2B mutant in which the 2B subdomain is removed[33,35,36], towards an autoinhibitory function possibly regulating spatiotemporal activity[34] . However, removal of 1B also prevents displacement of nucleoprotein complexes by the helicase[36] and *repΔ2B* is unable to complement for *Δrep in vivo*[35,36]. The importance of the 2B conformation on Rep activity agrees with its homologues UvrD in *E. coli* and PcrA in *Geobacillus stearothermophilus* in which 2B can rotate between 130° and 160° relative to the other three subdomains. In these homologues, the "closed" conformation confers a greater processivity and level of unwinding than in the "open" [9,11,25–29 23–28,37,38] form. Similarly, Rep that has been crosslinked in the "closed" state acts as a "super helicase" in terms of processivity and force generated[26]. Although wild type monomers of superfamily 1A helicases are unable to unwind significant lengths of dsDNA, this activity is activated when multiple helicases act on the same DNA molecule[28,34,39–42] or when activated by another co-factor as in the case of UvrD and MutL or PcrA and RepD[43–45] and generally correlates with the 2B subdomain being pushed towards a more "closed" state. Recently, the application of single-molecule Förster resonance energy transfer (smFRET) experiments with 32 ms time-resolution, suggested that the 2B subdomain in immobilized UvrD helicases populates as least four interconverting rotational conformational states on its DNA substrate[28] . However, whether this feature is a characteristic of other 1A superfamily helicases, and if this structural heterogeneity exists in the absence of the DNA substrate, remains unclear.

Here, we have used time correlated single photon counting, fluorescence correlation spectroscopy, time-resolved smFRET, Anti-Brownian ELectrokinetic (ABEL) trapping and total internal reflection

fluorescence (TIRF) microscopy and integrated these high-precision data with all-atom molecular dynamics (MD) simulations, to study freely-diffusing and immobilized Rep molecules. We find that in the absence of its DNA binding partner, Rep exhibits at least four different metastable structural conformations, two of which correspond to the open and closed states identified from previous static investigations using X-ray crystallography of the DNA-protein complex[24]. We observe that changing the ionic strength from high to low biases the conformations from open to closed, however, we find that multiple reversible transitional pathways exist that involve at least two additional intermediate states involving significant movement of all four of the protein's subdomains as opposed to just a single simple rotation of 2B around 2A as has been suggested for UvrD during its translocation activity on DNA[25]. We find that five out of the 12 transitions which are in principle possible between any pair of these four states show a strong rate dependence on sodium chloride concentration. Simulations indicate that, upon DNA binding, the extent of Rep conformational dynamics is significantly reduced. Our findings point towards an exploratory strategy that is more complex than previously reported for UvrD. This more granular transitional pathway may enable superfamily 1A helicases, such as Rep and UvrD, to reversibly trial a range of structural states steered by ionicity of the surrounding aqueous solvent prior to locking-in on a more rigid conformation upon binding to DNA. Our observations pave the way for exploring conformational dynamics of helicases in response to accessory factors, and for further elucidating the mechanism by which nucleoprotein complexes are displaced to ensure genomic integrity, and may reveal a more general strategy for ionic regulation of enzyme-substrate binding where multiple conformational states of the enzyme are involved.

Results

Increasing ionic strength results in a more open Rep conformation

To investigate the conformational states of free Rep we generated a Rep Δcysteine mutant from the wild type gene and used this to create a Rep two cysteine mutant by mutating alanine residues 97 and 473 located on subdomains 1B and 2B respectively to non-native cysteine (Methods). This mutant was cloned into *E. coli* and characterised for functionality using complementation assays (Supplementary Fig. 1, Supplementary Table 1 and Methods)[46]. Following overexpression and purification, we incorporated fluorescent tags onto the cysteine sites using a preliminary range of candidate FRET donor/acceptor pairs and concentrations, with donor Alexa Fluor 546 and acceptor Alexa Fluor 647 yielding the best response for the maximum number of FRET-active species (Supplementary Fig. 2) to report on conformational closing as suggested from static X-ray crystallographic structures of open and closed states[24] (Fig. 1a). Using fluorescence spectrophotometry, we estimate that most Rep molecules contained one molecule of each dye, with no likely preference for either dye to either cysteine residue. We confirmed that this labelled mutant retained helicase function by measuring its ability to unwind DNA *in vitro* relative to the unlabelled wild type protein. (Fig. 1b,c). Phosphoradioimaging of radiolabelled DNA (Methods) indicates that labelled A97CA473C Rep unwound 74.6 ± 4.7% of forked DNA compared to wild type 71.7 ± 1.3%, consistent with retention of activity.

We first investigated Rep in solution with ensemble FRET spectroscopy through excitation of the Alexa Fluor 546 donor tag using time correlated single-photon counting (TCSPC) to measure the fluorescence lifetime and anisotropy. FRET between the donor and acceptor is inversely proportional to the 6th power of inter-dye distance, such that an increase leads to donor quenching manifested as a distance-dependent shortening of its fluorescence lifetime[47]. We measured the lifetime in low and high salt buffers corresponding to 10 mM and 500 mM NaCl and determined the lifetimes from fitting of the fluorescence intensity decay profiles with a sum of multiple exponential components (Methods). For low salt, the amplitude-weighted average fluorescence lifetime (τ_{av}) was ~1.5 ns, whereas for high salt τ_{av} increased more than 2-fold to ~3.5 ns, indicative of a substantial increase in

the inter-dye distance corresponding to a more open conformation. Based on donor dye-only controls under identical conditions (Supplementary Fig. 3a-d), this increase in lifetime was not due to photophysical artifacts associated with the free dye. However, when similar experiments were performed with Rep labelled only with Alexa Fluor 546 at 473C or at 97C, we measured a similar initial ensemble lifetime and comparable increase with salt addition implying the donor also displays relatively small but measurable fluorescence shifts which could be attributed to conformational flexibility, possibly via a mechanism involving protein-induced fluorescence enhancements (PIFE). The smallest number of fitting parameters that accounts for the fluorescence lifetime of Alexa Fluor 546 on Rep corresponds to a three-component model (Supplementary Table 2), hinting at the presence of multiple different Rep conformations, though with other contributory factors such as ground-state variability of the dyes and intrinsic fluctuations in molecular structure potentially also contributing to fluorescence decay heterogeneity. Anisotropy measurements suggested negligible impairment to dye rotation when bound to Rep.

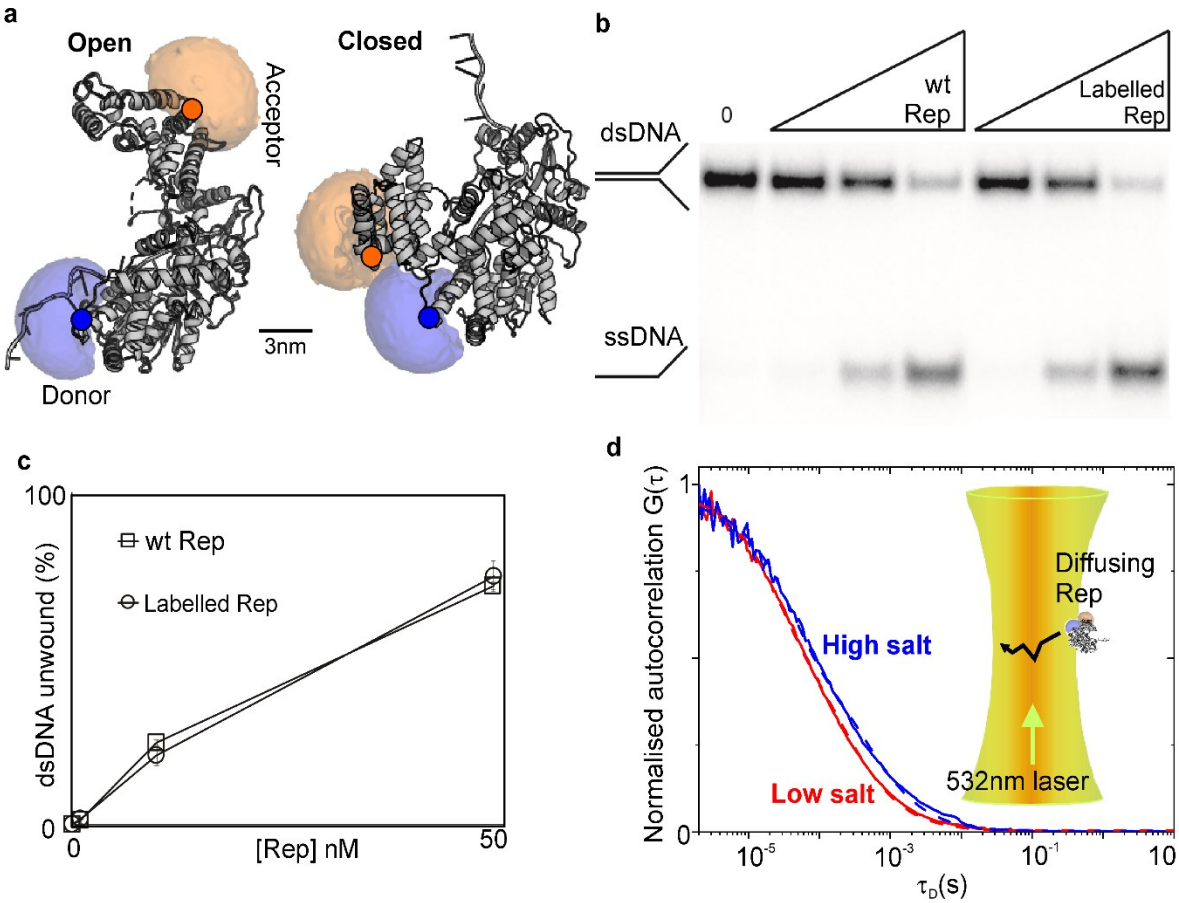

Figure 1. FCS indicates significant ionic strength dependence on Rep conformation. a. Crystal structures of Rep occupy both open (left) and closed (right) conformations, suggesting rotation of 2B subdomain. Donor and acceptor dyes here labelled at A97C and A473C (orange and blue filled circles respectively) via a maleimide linkage represented by accessible volumes shown (donor: light blue, acceptor: light orange, note with our labelling protocol it is also possible to have donor at A473C and acceptor at A97C), PDB ID: 1UAA. **b.** 10% TBE polyacrylamide gel showing unwinding of radiolabelled forked substrate, to produce ssDNA. Wild type and labelled Rep concentrations were 1, 10 and 50 nM, plotted in **c.** The **d.** FCS normalised autocorrelation, G, plotted as a function of translation diffusion time τ_D, of 30 nM Rep at low (red) and high (blue) salt concentrations, heuristic fits (dashed) are also indicated.

We used fluorescence correlation spectroscopy (FCS), a well-characterized single-molecule tool that can be applied to probe the mobility of ensembles of solvated biomolecules[48–52] as a preliminary method to probe the effects of changing ionic strength on Rep's structural conformation by measuring its effect on molecular mobility. Although FCS has been used to investigate helicase-catalysed unwinding[51] and the interactions between the RecQ helicase, Rep and short DNA substrates[52], it has not been previously applied to Rep dynamics. We used donor excitation of solutions of dual-labelled Rep in high and low salt buffer to determine normalized fluorescence autocorrelation curves (Methods). Since the hydrodynamic radius of diffusing molecules scales approximately with the cube route of molecular weight, FCS is poor at discriminating different types of molecules unless the difference in molecular weight is typically at least a factor of four[53]. With this caveat, we found that increasing ionic strength even so resulted in a measurable shift to longer characteristic translational diffusion times through the confocal volume (typically a few hundred to a few thousand µs) (Fig. 1d) indicating a higher frictional drag consistent with a larger effective hydrodynamic radius associated with Rep. Taken together with the ensemble FRET spectroscopy measurements, these data suggest substantial conformational rearrangement for Rep from closed to more open and less compact conformations in changing from low to high salt.

Single-molecule FRET reveals free Rep exhibits at least four states

Although FCS is capable of detecting single-molecule fluorescence bursts, both it and ensemble FRET spectroscopy report on population properties that potentially mask molecular heterogeneity[54–57]. To investigate the dynamics of individual Rep molecules, we first tried dual-colour total internal reflection fluorescence (TIRF) and variable-angle epifluorescence microscopy (VAEM)[58] to visualise wide fields of view of labelled Rep molecules immobilised to a coverslip surface (Methods). However, although donor- and acceptor-only excitation generated similar numbers of fluorescent foci per field of view, the yield of detectable anti-correlated behaviour from colocalised donor/acceptor foci was poor due to surface-induced dye photoblinking (Supplementary Fig. 4) as well as proteins labelled with either two donors or two acceptors. To mitigate against surface effects, we then tried smFRET detection using confocal laser excitation ~20 µm deep into the solution on freely diffusing Rep (Fig. 2 and Methods). As for the earlier FCS measurements, diffusion of labelled Rep molecules could be monitored as fluorescence emission during the typically ~millisecond transit time through a confocal laser excitation volume. However, here we used rapid µs alternating laser excitation (ALEX) for excitation of the donor and acceptor separately that could be reconciled into coincident millisecond timescale bursts of fluorescence in donor and acceptor channels (Supplementary Fig. 5).

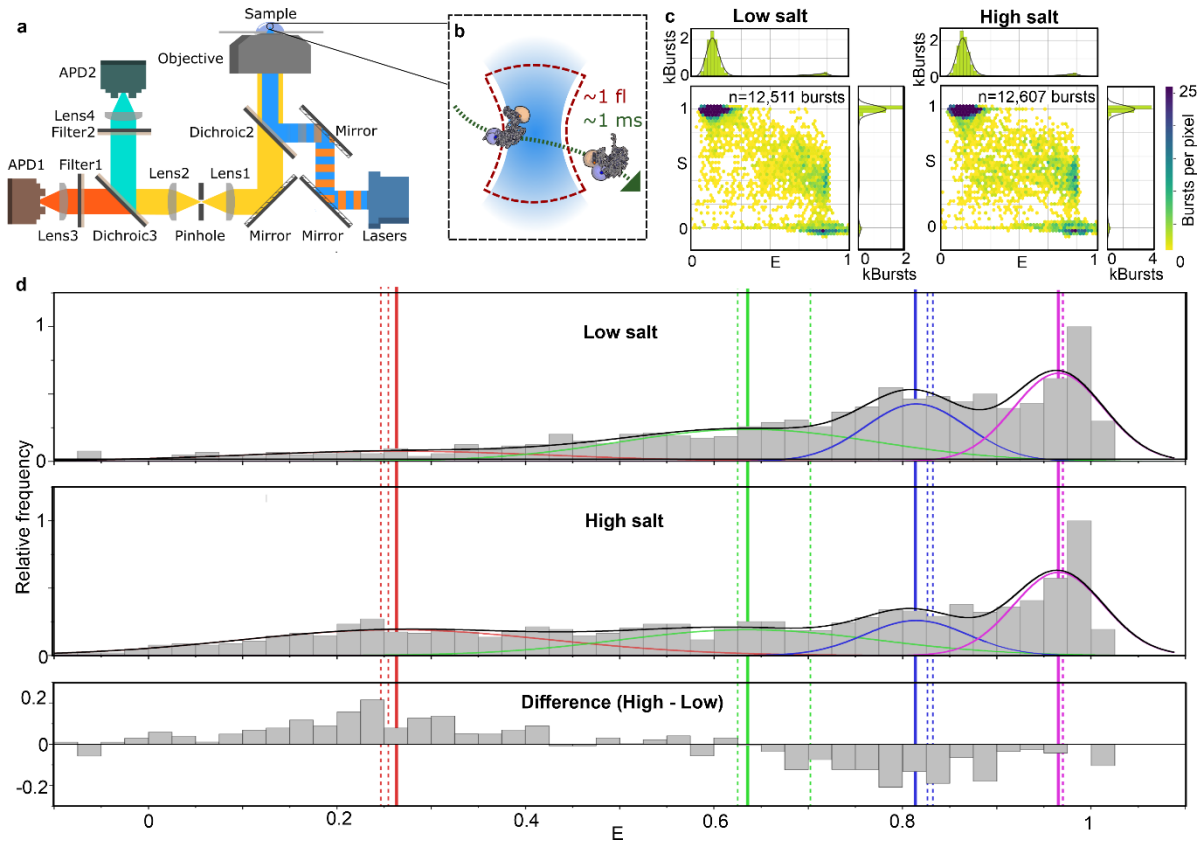

Figure 2. Single-molecule FRET using confocal microscopy reveals a multiplicity of ionicity dependent Rep structural conformations. a. Schematic of the smFRET instrumentation with **b.** a labelled Rep (not to scale) diffusing through the confocal volume. **c**. E-S plots for Rep in low salt 10 mM (number of detected burst events n=12,511) and high salt 500 mM NaCl (n=12,607). Note that the number of burst events is higher than those used in the 1D histograms as they include the singly and doubly labelled populations. **d.** Distribution of FRET efficiency for DA-Rep population using confocal alternating laser excitation smFRET in (top panel) low and (middle panel) high salt, four Gaussian fit (red – state S1, green – S2, blue – S3, magenta – S4, and summed total in black), and (bottom panel) difference between low and high salt distributions. Centre of Gaussians indicated in vertical solid lines, dashed lines represent FRET efficiency populations recovered from H2MM with accurate FRET correction parameters applied (Methods). Taken using number of bursts = 1,021 and 1,013 from high and low salt respectively, from 7 different samples.

Photon counts from each channel were converted into estimates of FRET efficiency E and relative stoichiometry S of donor and acceptor labelling on a molecule-by-molecule basis (Fig. 2b). We detected a greater proportion of D-only labelled Rep compared to DA- Rep, of ~90, illustrating the value of using ALEX in conjunction with the E-S relation to determine the DA-Rep population for FRET analysis. Analysing the DA-Rep population (i.e. those which are labelled with both a donor and acceptor and thus capable of undergoing FRET) in low salt indicates a broad distribution of FRET efficiency values across the full range of 0-1, but with a bias towards high values; in high salt, the FRET efficiency distribution was similarly broad but with a bias towards lower values (Fig. 2c). Using a multiple Gaussian fit we could account for these two distributions with four FRET states S1-S4 of mean values 0.26 ± 0.06, 0.64 ± 0.07, 0.81 ± 0.01 and 0.97 ± 0.01 (Supplementary Table 3). The proportion of Rep in each state by area of each respective Gaussian is approximately 13:33:23:32% and 31:26:13:30% at low and high salt respectively for S1:S2:S3:S4. Analysing the differences between the

low and high salt distributions indicated that largest discrepancies were centred around the low FRET state of 0.26 ± 0.06 and the higher FRET state of 0.81 ± 0.01.

In order to investigate the lifetime of Rep conformations beyond the typical millisecond timescale transit time through a confocal volume, we were able to monitor single free molecules for extended time durations up to three orders of magnitude greater using an Anti-Brownian ELectrokinetic (ABEL) trap[59] even for high salt conditions[60], optimising preliminary ABEL measurements on Rep made in low salt non-physiological buffers[46]. The ABEL trap maintains the position of a trapped single molecule in a confocal volume in free solution within a microfluidic cell, away from the perturbative effects of surfaces and able to rotate freely similar to confocal smFRET, but using advanced closed-loop electrokinetic feedback to dynamically recentre a diffusing tracked molecule between four microfabricated electrodes. The trap location is conjugate to sensitive detectors that can record simultaneous sample brightness, emission colour, lifetime, and anisotropy, among other spectroscopic and physical variables[61–64]. A diagram of the ABEL trap and examples of raw data from individual trapped DA-Rep molecules are shown in Fig. 3a,b. In the raw trapping traces, anti-correlated brightness of donor and acceptor are evident, along with transitions among multiple FRET states. These data show that both high and low FRET states are accessed in both salt conditions, but that extended high and low FRET states are more common in low and high salt, respectively.

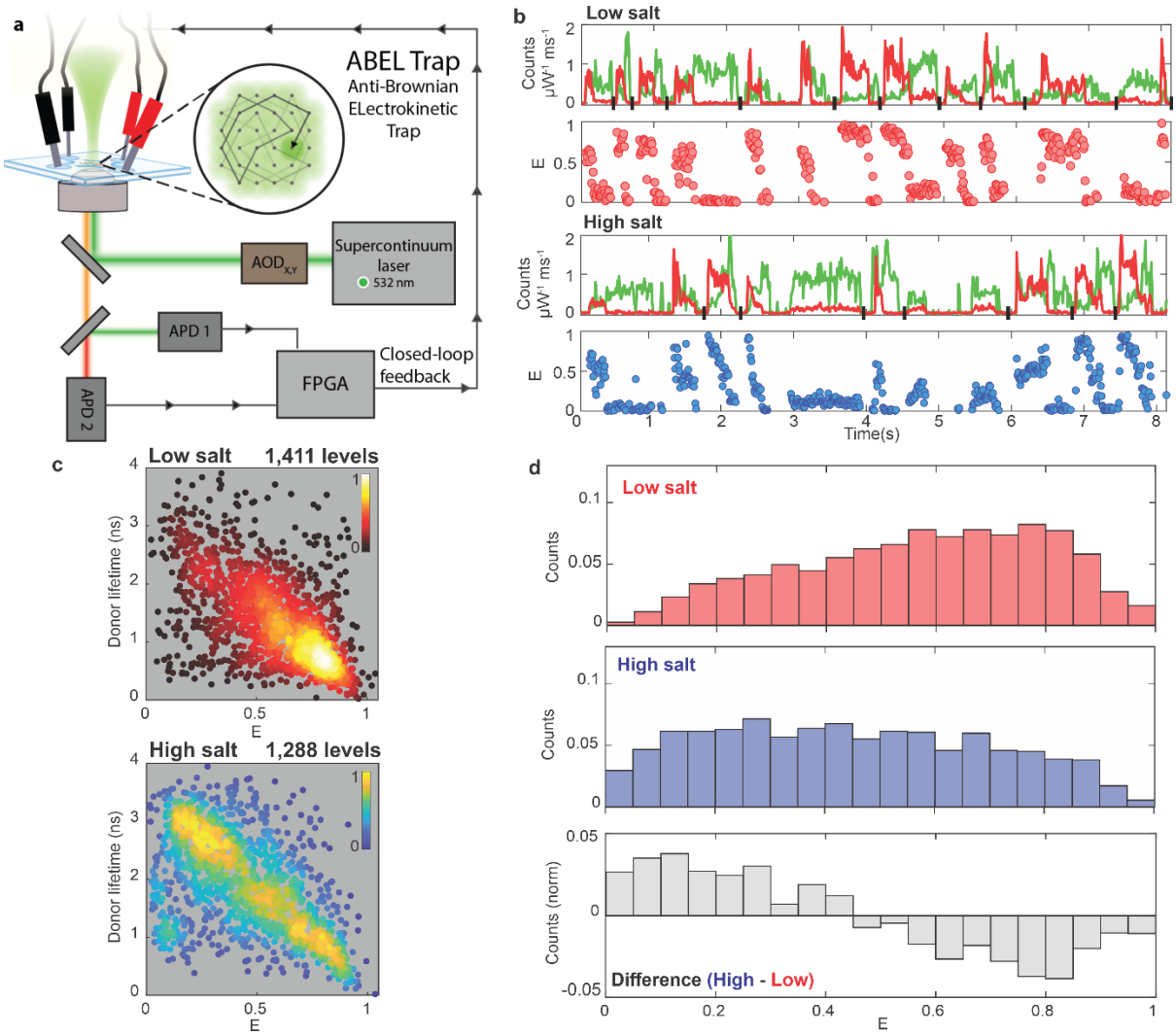

Figure 3. Lifetime and FRET distributions of Rep measured in the ABEL trap are salt-dependent. a Diagram of ABEL trap setup as used in this work. Pulsed, linearly polarized 532 nm wavelength excitation is scanned by AODs across a microfluidic trapping region; photon-by-photon readout is used

to determine closed-loop voltage feedback to X and Y electrodes to counteract Brownian motion. **b.** Raw ABEL trap brightness data (top) and FRET values (bottom) for several sample events (separated by black vertical slashes on abscissa) show transitions occurring on broad millisecond to second timescales in both low salt, here using 25 mM (top), and high salt 500 mM NaCl (bottom). Multiple distinct levels are observed, along with dye photophysics. Brightness traces are normalized by excitation power and bin size and corrected for different detection efficiency due to detection filters in each channel. **c.** Scatter plot of measured fluorescence lifetime vs. FRET efficiency for DA-Rep population using 532 nm wavelength excitation in low (red; top) and high salt (blue; bottom); each point represents one brightness level, colouration reflects local relative density. **d.** Histogram of FRET efficiency in low (red) and high (blue) salt, with difference in normalized FRET efficiency histograms for high minus low salt data indicated (grey).

Fig. 3c shows scatter plots of simultaneously measured fluorescence lifetime and FRET efficiency, where each point represents a single level as identified by change point analysis (see Methods). Each scatter plot contains data from >1,000 levels where both donor and acceptor were photoactive. The colour of each point reports the relative local density of points. These data confirm a correlated shift from high FRET efficiency and short donor lifetimes in low salt to a lower, broader FRET efficiency distribution in high salt. One-dimensional histograms showing the FRET efficiencies from this ABEL trap data are shown in Fig. 3d, with equivalent difference histograms, which show a similar pattern to the result obtained by confocal microscopy with ALEX as described above, but demonstrating that some FRET states can persist for an extended timescale of ~seconds.

Rep undergoes dynamic transitions over a milliseconds-seconds

Preliminary Burst Variance Analysis[65] (BVA) on confocal laser excitation smFRET indicated that the variance of the fluorescence intensity during diffusional transit was broader than that expected if just a single FRET state was present, suggesting transitions between FRET states over a millisecond timescale. To investigate this quantitatively, we applied H2MM analysis. The H2MM algorithm is a specialised hidden Markov model designed for analysing single-molecule fluorescence data. It enables characterisation of biomolecular conformations and kinetics by processing the sequential photon-by-photon data from within bursts utilising photon arrival times to determine the dynamic states and transitions of molecules through maximum likelihood estimation[66] (Methods). We used these predictions to infer the likely number of states present and to determine optimised values based on a Bayesian information criterion for the associated kinetic rate constants. Results from the H2MM analysis indicate a good agreement for both low and high salt data for a four-state model based on a modified Bayes information criterion (BIC')[67] (Fig. 4 and Supplementary Table 4). The analysis suggests that two states (identified as S1 and S4) exhibit FRET values which are close to that expected for the known crystal structures for closed and open conformations of Rep. Fig. 4a and 4b indicate the estimated rate constants between all four states from this analysis, while Fig. 4c depicts this as the change to each transition between low and high salt. This analysis indicates that the kinetics between S1 and S4 primarily involve intermediate transitions between metastable states S2 and S3, with transitions between S2-S4 being relatively unlikely. Also, changing the ionicity from low to high results in significant changes to the S1-S3 transitions (both forward and reverse rates are significantly *decreased* in high salt) and the S2-S3 transitions (both the forward and reverse rates are significantly *increased* in high salt).

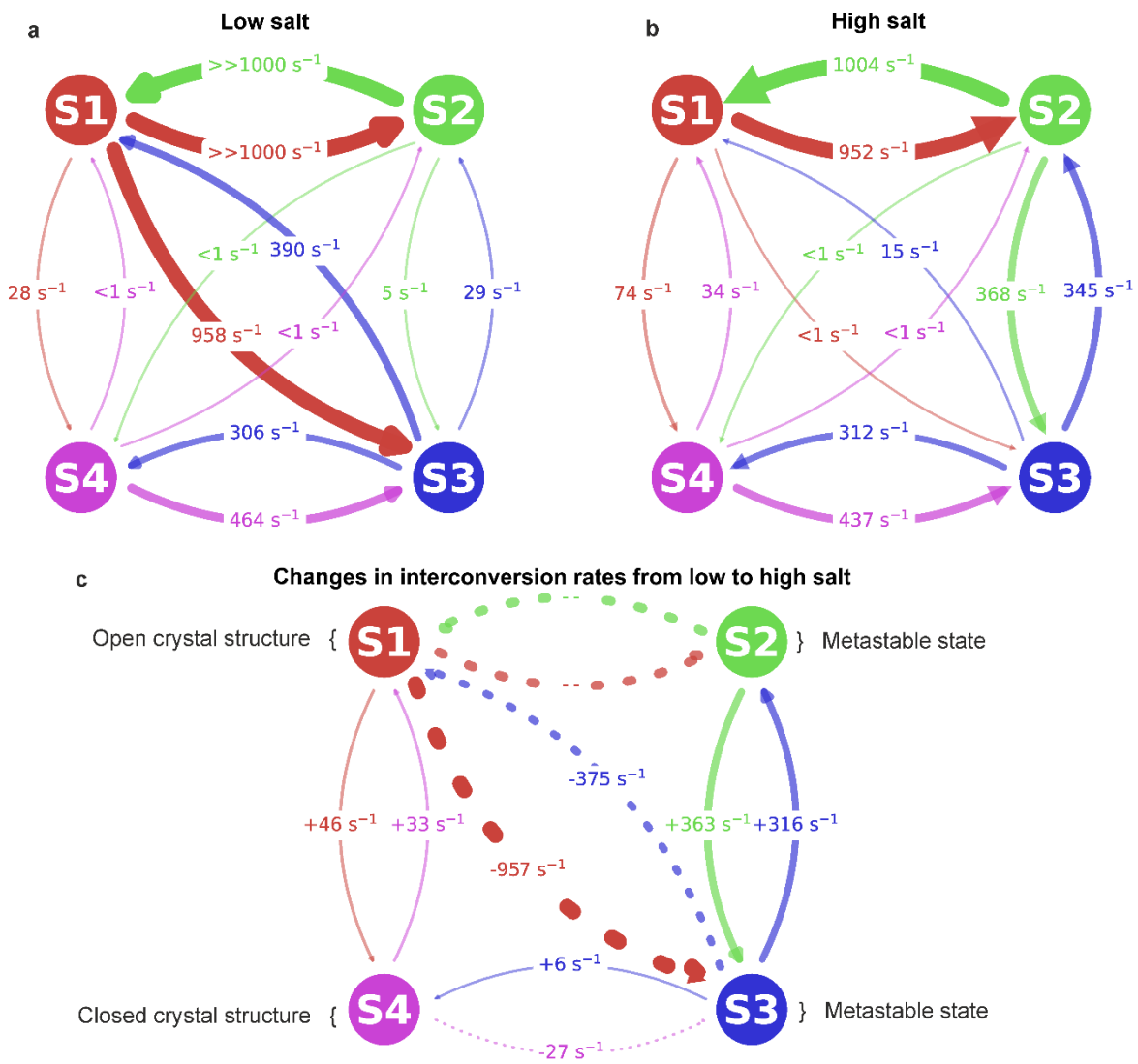

Figure 4. Rep undergoes millisecond timescale FRET transitions between four states. H2MM analysis on confocal excitation smFRET data at **a.** low salt 10 mM and **b.** high salt 500 mM NaCl, showing the estimated transitional rate constants between the four identified FRET states (same colour codes as the states indicated from four Gaussian fit in Fig. 2), with optimised rate constants shown between these. **c.** The change in rate constants between low and high salt. Note, since the S1-S2 transitions in low salt typically occur faster than the nominal fastest ALEX sampling time of 100 µs we cannot quantify their absolute change in going from a low to a high salt environment, though we can measure that the forward and reverse rates in high salt although at the limit of our temporal resolution are both likely lower than in low salt. Similarly, since the S2-S4 transitions are relatively rare at <1 s^{-1}, the absolute changes in interconversion rates due to changes in ionicity cannot be quantified.

Metastable FRET states are associated with distinct Rep structures whose conformational freedom is limited by DNA binding

To establish the mechanistic features of the Rep conformational transitions, we used atomistic molecular dynamics (MD) simulations to Angstrom-level resolution (Methods). The starting structures were derived from the crystal structure of Rep (PDB ID 1UAA[24]) with (i) open and (ii) closed conformations initially deleting the DNA. For capturing large conformational transitions under low and high salt, we used 50 ns windows in implicit solvent, running four replicates for each of the four

conditions. Then, time-dependent conformational changes were evaluated (Fig. 5a and Supplementary Movies 1-4).

In particular, the straight-line distance between residues 97 and 473 was evaluated to test for transient intermediate conformations between the labelling sites chosen for smFRET measurements. For simulations starting with the open conformation, we observed significant variability in this distance (approximately from 5 to 11 nm) regardless of ionic concentrations, which is consistent with Rep maintaining an open structure (Supplementary Fig. 6). When the simulations started from the closed conformations, this distance stayed in the closed state (around 3 nm) for all low-salt replicas, while it gradually increased from 3.5 nm to 8 nm, consistent with a progressive opening of the structure from the closed form for some of high-salt replicas (Supplementary Fig. 6). A key observation is thus that switching between states was observed, consistent with a multi-state conformational landscape as opposed to a simple two-state system identified by the crystal structure.

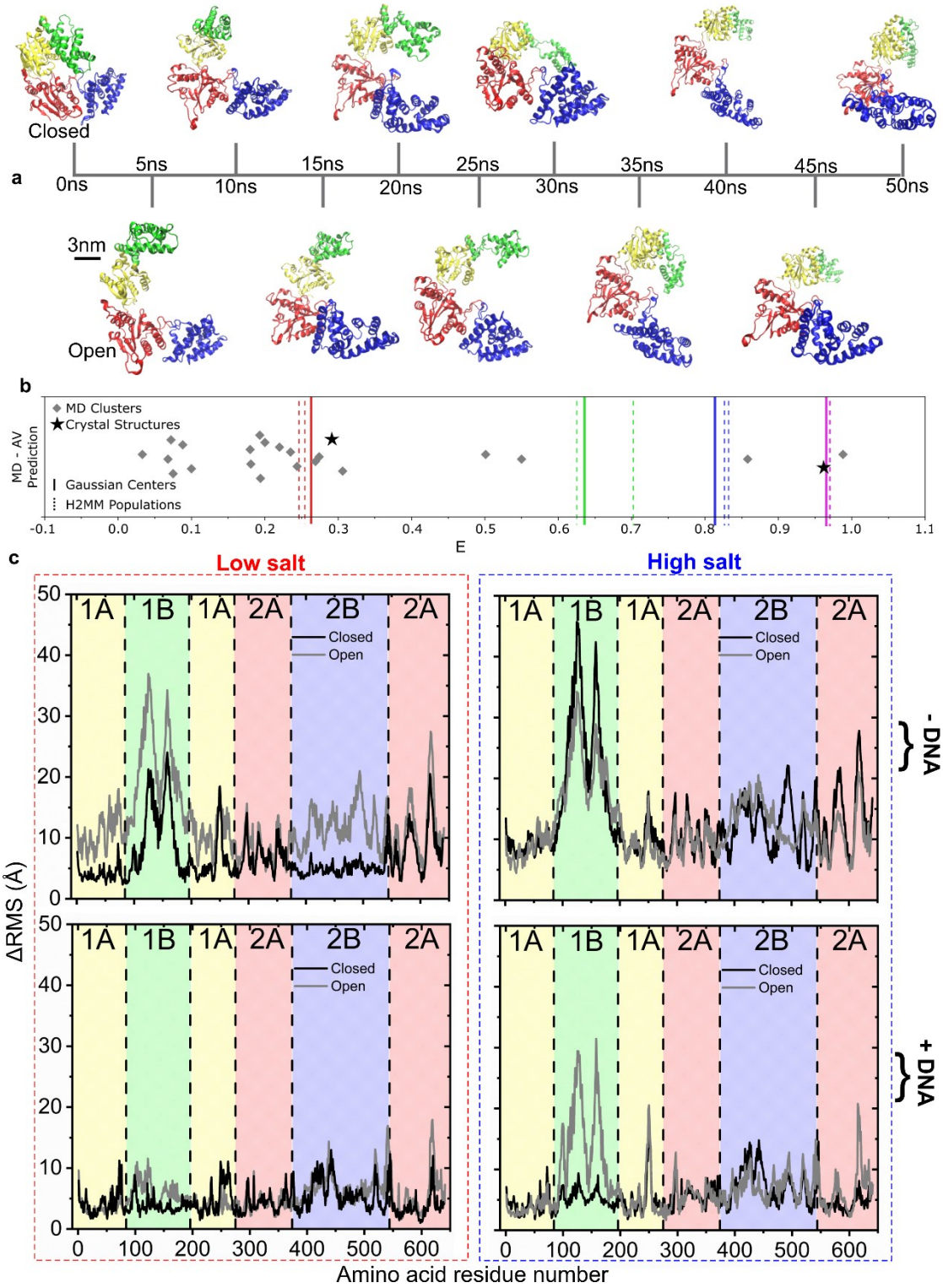

Figure 5. MD simulations on Rep indicate multiple structures whose metastability is dependent on ionic strength. a. Snapshots of predicted Rep structures over a 50 ns simulation implicit solvent in 10 mM monovalent salt NaCl, starting from either the fully open (top panel) or closed (bottom panel) crystal structures. **b.** Cluster analysis on MDS predicted structures follow by FRET AV analysis indicate reasonable agreement to the experimental observations of four different Rep FRET states, centre of Gaussians from Fig. 2 overlaid in vertical solid lines with dashed lines the values from H2MM analysis (Methods). **c**. Root-mean-squared deviation (RMSD) per residues averaged over all replicas starting from the open and closed crystal structures suggests significant flexibility in the 1B domain in low salt

(10 mM) and high salt (500 mM) with no DNA bound to Rep (upper panels). Binding of DNA to Rep (lower panels) results in a significant overall reduction in RMSD apart from the 1B domain in the open Rep conformation at high salt.

We performed cluster analysis to identify the most likely occurring structures from the full suite of simulated data at low and high salt, then FRET accessible volume (AV) analysis[68] on these most probable structures to predict the measured values for FRET efficiency using the known position of the dye tags and incorporating realistic flexibility from the linkers used (Methods). Comparing these predicted FRET efficiency values with those optimised from H2MM analysis and with the mean values for each respective Gaussian from the multi-Gaussian fit analysis indicated good agreement towards a four-state model comprising an open and closed state S1 and S4 which correspond to the known crystal structures at low and high FRET efficiency values respectively, plus two intermediate states S2 and S3 (Fig. 5b). Similarly, estimating the radius of gyration from the MD simulation structures indicated a distribution between approximately 25-45 nm across the different simulation conditions, broadly comparable to the experimental measurements from FCS (Supplementary Fig. 7a-d and Methods), which could be fitted with up to four Gaussians, consistent with a heterogeneity of molecular conformational states[69].

Root mean square deviation (RMSD) analysis of all collected MD trajectories (Fig. 5c) indicated high flexibility in the 1B domain for both low and high salt conditions in the absence of any bound DNA. In high salt, the open and closed forms of Rep display striking similarities, but at low salt we see a decrease in flexibility of the 1B and 2B domain in the closed conformation. The high flexibility of the 2B domain can also be seen from population radar plots from these simulations of the rotation angle of the hinge region of Rep which connect the 2A and 2B domains[23–28,37,38] (Supplementary Fig. 7e,f and Methods), possibly mediated by the charged interface between an alpha helix in the 1A subdomain (position R195-K207) and an alpha helix in the 1B subdomain (position L87- L99). Apart from the high 1B flexibility in the open conformation also being seen at high salt when DNA is bound, DNA binding results in a substantial reduction of flexibility for all other domains at both low and high salt as well as for the 1B domain in the closed conformation.

Discussion

Here we investigate the structural plasticity of single enzyme molecules in the absence of any bound substrate. We explored this phenomenon in Rep, an example of a prokaryotic accessory helicase, of the 1A superfamily, known to have structural and oligomeric complexity when bound to their DNA substrate[23–25,34,70]. Rep helicase is monomeric in solution and can bind to ssDNA as a monomer and is able to translocate in a highly processive and rapid fashion coupled to ATP hydrolysis, however monomeric Rep is poor at unwinding dsDNA that it encounters with any interactions resulting in unwinding rare[71]. The efficacy of Rep as a helicase is greatly improved when Rep monomers act in concert or as dimers[72–74] or when activated by cofactors such as PriC[75]. Two monomers of Rep have been crystallised on the same piece of ssDNA[24]and show that upon binding to ssDNA Rep can occupy at least two conformational states, that vary by virtue of the 2B domain having rotated through ~130°. Studies have shown that the 2B subdomain plays an autoinhibitory role in the processive unwinding of DNA[33,71] and is essential for removing nucleoprotein blocks ahead of the replication fork[36] . The 2B subdomain of Rep is also able to move with regard to the rest of the protein whilst translocating along ssDNA[23] and adopts a closed conformation when approaching a duplex region of DNA.

Whilst recent studies on the homologous UvrD helicase have shown that its 2B subdomain can occupy four distinct conformations two of which map with the open and closed crystal structures of UvrD with the other two states being metastable intermediates[25,28]. However, the four

conformational states of UvrD were visualised by surface immobilising UvrD onto a coverslip for smTIRF microscopy and it is unclear how this surface tethering affects the conformational landscape of a protein. All of this taken together motivated the need to investigate the conformational landscape that Rep can occupy and ascertain whether having a 2B subdomain that occupies a number of conformational states might be a common feature of all superfamily 1A helicases.

Initially, we used lifetime measurements of doubly labelled Rep indicating significant ensemble-level conformational heterogeneity. Using low and high NaCl concentrations we were also able to tune the fluorescent lifetimes of donor dyes on the protein, donor lifetimes more than doubling under high NaCl in agreement with the effect of changing NaCl concentrations on the homologous UvrD protein[28] and suggesting an opening of the 2B subdomain with increasing NaCl.

We first attempted to investigate the different Rep conformations at a single-molecule level using TIRF microscopy, in which Rep was fixed to a microscope coverslip. TIRF microscopy proved useful in that it substantiated our premise that free Rep is conformationally dynamic involving more than one structural state. However, these data were dominated by FRET traces that either showed large aggregates of Rep protein with exponential like decay or traces that contained non-correlated FRET fluctuations. We attributed these observed non-correlated FRET fluctuations to a large degree of dye photoblinking caused by being in close proximity to the coverslip surface, which motivated our attempts to visualise Rep conformational dynamics without having to surface-immobilise the protein.

Using single-molecule confocal microscopy in solution which removed confounding effects of surface immobilisation, we are able to identify several distinct FRET states that Rep can occupy in the absence of any cofactors. Here, the idealized FRET trajectories were reconstructed in a probabilistic manner using an information minimisation criterion, enabling four states, rather than the two observed by crystallography, to be revealed. Under the solution conditions tested, the FRET states were observed to rapidly interconvert freely in solution, reflecting dynamic conformational heterogeneity in the helicase, even in the absence of any cofactors, further substantiated by utilising ABEL trapping which demonstrated relatively long-lived FRET states up to a timescale of ~seconds. Although our measurements are consistent with a model developed previously for UvrD that the 2B subdomain rotates about the hinge region connecting it with the 2A subdomain[25], our findings from Rep, which has a large degree of structural and sequence homology with UvrD, suggest that this transition is unlikely to be a simple single arm-swing mechanism.

We compared the interconversion rate constants from our measurements on Rep with those reported for similar states identified in UvrD using smFRET measurements from dye pairs positioned at equivalent locations[28] but using a much slower sampling time of 32 ms compared to that of our detection system which was 2-3 orders of magnitude smaller (Supplementary Table 5). These comparisons clearly show that faster sampling reveals hitherto hidden transient intermediate states, whose interconversion kinetics in several cases are strongly dependent on salt concentration. Specifically, in changing the NaCl concentration from low (10 mM) to high (500 mM) indicated the largest changes in interconversion rate constants for those involving the metastates S2 and S3; for S2 to S3 with factor x74 *increase* and for transition S1 to S3 with a x1000 factor *decrease*. These findings offer new insight into how the local ionicity of the surrounding water solvent can in effect tune the structural landscape of accessory helicases.

MD simulations run at high and low salt concentrations, and using starting points of the open and closed crystal structures for Rep, followed by FRET AV modelling to predict the associated level of FRET efficiency with different consensus structural predictions, showed good agreement to experimental measurements of the four measured FRET states. Structural heterogeneity was also measured from associated radius of gyration predictions and conformational multicity associated with rotation about the hinge region joining the 2A and 2B subdomains. The absence of any bound

DNA substrate from these simulated structures resulted in substantial high RMSD values compared to simulations involving DNA bound to Rep, consistent with increased structural plasticity in the absence of bound substrate, which locks into less flexible structures upon DNA binding. In particular, we found that the closed structure was much more stable than the open, and only the closed structure in high salt moves away from the open conformation within the simulation period.

We find that dynamically transitioning between metastable states of the free Rep protein results in a bias towards either fully open or closed conformations depending on ionic strength, however, the proportion of Rep in open:closed S1:S4 states at low and high salt is not 0:100% and 100:0% respectively, but closer to approximately 12:33:23:32% and 31:26:13:30% for S1:S2:S3:S4. This distribution of conformational states could offer a bet-hedging strategy to ensure that there are still reasonable numbers of free Rep adopting conformations which are either close to the fully open or closed DNA-bound conformations at any point in time thereby facilitating responsive and rapid adaptation between these two states if required physiologically. Having a structural conformation for free Rep that are close to those of the DNA-bound states may also enable a more rapid conformational response upon DNA binding and unbinding during the enzyme's catalytic activity of aiding nucleoprotein block removal on DNA. Our molecular dynamics simulations suggest that the IB subdomain region shows the highest sensitivity to ionic strength which may be meditated by charged residues in the vicinity of the hinge involving the charged interface between the 1A subdomain (R195-K207) and the 1B subdomain (position L87- L99). Such structural plasticity dependent on ionic strength is consistent with earlier experimental observations of accessory helicase UvrD and therefore may suggest present a more general feature of this 1A helicase superclass. Future work to test this hypothesis further could involve establishing the mechanistic role of cofactors such as the DNA substrate and of bound ATP, and also to determine whether the closed state of Rep is needed to promote interactions or confer activity.

Methods

Labelling Rep. *rep* ΔCys A97CA473C mutant was purchased as a gene synthesis product in a pEX-K plasmid backbone from Eurofins MWG in which native cysteines at positions 18, 43, 167, 178 and 617 were mutated to L, S, V, A and A respectively as per Rasnik *et al.*[76] as well as the A97C and A473C mutation. Rep ΔCys A97CA473C was cloned from this plasmid into pET21a via restriction sites NdeI and XhoI creating plasmid pJLH103. Rep ΔCys A97CA473C was subcloned from pJLH103 into pET14b using restriction sites NdeI and XhoI producing His-tagged Rep ΔCys A97CA473C, pJLH135. Rep ΔCys A97CA473C was also cloned into pBAD24 (Kan^r) plasmid for use in complementation assays. The *rep* ΔCys A97CA473C mutant was excised from pEX-K *rep* ΔCys A97CA473C plasmid via sequential NdeI restriction, klenow treatment followed by PstI restriction. This fragment was ligated into a pBAD24 plasmid that had been sequentially XmaI restricted, klonow treated and PstI restricted, generating pJLH121. Complementation of *Δrep ΔuvrD* lethality by pBAD constructs was performed as previously described[15]. Briefly, pBAD plasmids were transformed into N6524 and N6556. Both strains contained plasmid pAM403 required for maintenance of cell viability of *ΔrepΔuvrD* strains on rich media. pAM403 is highly unstable and lost by re-streaking on minimal media whilst maintaining selection of pBAD. Colonies were picked and grown in liquid minimal media to stationary phase before being serially diluted and spotted on LB agar containing kanamycin ± arabinose and plates photographed after 24 hours at 37°C. Assays were performed twice.

His-tagged Rep ΔCys A97CA473C was overexpressed from the relevant pET14b clones in HB222 (BL21AI*Δrep*). For overexpression, cells were grown to OD_{600} 0.7 shaking at 37°C, induction performed at 20°C with 0.2% arabinose for 3 hours. Cells were clarified, and flash frozen in 50 mM Tris pH 7.5 and 10% (w/v) sucrose. Clarified cells were thawed on ice and additions then made so that the suspension contained 50 mM Tris-Cl pH 8.4, 20 mM EDTA pH 8.0, 150 mM KCl and

0.2 mg ml^{-1} lysozyme. After 10 min incubation on ice, Brij-58 was added to 0.1% (v/v; final concentration) with a further 20 min incubation on ice. Supernatant was recovered by centrifugation (148,000 x g, 4°C for 60 minutes) and DNA precipitated by dropwise addition of polymin P to 0.075% (v/v; final concentration) with stirring (4°C, 10 minutes). The supernatant was recovered by centrifugation (30,000 x g, 20 minutes, 4°C) and solid ammonium sulphate added to 50% saturation with stirring (4°C, 10 minutes). The pellet was collected by centrifugation (30,000 x g, 20 minutes, 4°C) and stored overnight at 4°C. The protein pellet was resuspended in 20 mM Tris-HCl pH 7.9 and 5 mM imidazole until the conductivity matched 20 mM Tris-HCl pH 7.9 and 500 mM NaCl (buffer A) plus 5 mM imidazole. Rep ΔCys A97CA473C was purified by chromatography on a 5 ml His-trap FF column (GE Healthcare) using a 100 ml gradient of 5 mM to 1 M imidazole in buffer A. The conductivity of eluted protein from the His-trap column was adjusted to match buffer B (50 mM Tris pH 7.5 and 1 mM EDTA) plus 50 mM NaCl by dilution in buffer B. Rep proteins were then purified on a 3ml heparin-agarose column using a 60 ml linear gradient of 50 mM to 1M NaCl in buffer B. Peak fractions containing pure protein were collected and reduced for two hours at 4 °C by addition of 5 mM Tris(2-carboxyethyl)phosphine hydrochloride (TCEP). Ammonium sulphate was then added to the Rep solution under stirring to 70% saturation at 4 °C, stirring was allowed to continue for 10 minutes before centrifugation (18,000 x g, 20 mins at room temperature). The pellet was resuspended in degassed labelling buffer (100 mM sodium phosphate pH 7.3, 500 mM sodium chloride, 20% (v/v) glycerol).

We explored several FRET dye combinations initially using confocal smFRET (Supplementary Fig. 2) before settling on Alexa Fluor™ 546 and Alexa Fluor™ 647 since they indicated highest levels of 1:1 donor:acceptor labelling and minimal non-physiological aggregation. For optimised incubation conditions, Alexa Fluor™ 546 C_5 Maleimide and Alexa Fluor™ 647 C_2 Maleimide were dissolved in anhydrous DMSO and mixed in equimolar amounts; dyes were added to the reduced Rep sample in a fivefold molar excess. The mixture was rocked at room temperature 30 minutes before addition of 2-mercaptoethanol to 10 mM to quench the labelling reaction and left to rock for 10 minutes. The sample was loaded onto a 1 mL His-Trap FF crude column (GE Healthcare) equilibrated with buffer A (20 mM Tris-HCl pH 7.9 and 500 mM NaCl, 20% (v/v) glycerol) + 5 mM imidazole. The column was washed with 20 mL buffer A + 5 mM imidazole before being developed with a 20 mL gradient of buffer A + 5 mM imidazole to buffer A + 500 mM imidazole. Peak fractions containing labelled Rep were pooled aliquoted and stored at -80 °C in 20 mM Tris-HCl pH 7.9 and 500 mM NaCl, 30% (v/v) glycerol. Protein concentrations were measured using Bradford's assay and, in combination with individual dye concentrations as determined by absorbance measurements using NanoDrop™ 2000, labelling efficiencies and stoichiometries could be determined.

Helicase activity assays. Unwinding of streptavidin-bound forks was assayed using a substrate made by annealing oligonucleotides PM187B20:

(5' GTCGGATCCTCTAGACAGC(biodT)CCATGATCACTGGCACTGGTAGAATTCGGC) and PM188B34 (5' AACGTCATAGACGATTACATTGCTACATGGAGC(biodT)GTCTAGAGGATCCGAC). PM187B20 was 5' end labelled using 1 μL T4 polynucleotide Kinase (NEB, M0201S), 1 μL γ ^{32}P ATP (Perkin Elmer, NEG502Z250UC), 1 μg PM187B20 in a 20 μL reaction volume of 1x T4 PNK buffer (NEB, M0201S) for 60 minutes at 37 °C before heat inactivation at 65°C for 25 minutes. The reaction volume was brought up to 50 μL by addition of MilliQ water and run through a Micro Bio-Spin™ 6 column (Biorad, 7326221) as per manufacturer instructions to purify radiolabelled oligonucleotide from free γ ^{32}P ATP. For annealing, labelled PM187B20 was mixed with 3 μg of PM188B34 in 100 μL final volume of 1x sodium sodium citrate, heated to 95 °C five minutes before slowly cooling to room temperature. Annealed DNA was purified from a 10% Acrylamide TBE gel. Reactions were performed in final volumes of 10 μL in 50 mM Tris pH 7.5, 10 mM magnesium chloride, 2 mM ATP, 0.2 mg ml^{-1} BSA, 1 μM streptavidin and 1 nM DNA substrate. The reaction mixture was pre-incubated at 37°C for five minutes, then Rep helicase was added, and incubation continued at 37°C for 10 minutes.

Reactions were stopped with 2.5 µl of 2.5% SDS, 200 mM EDTA and 10 mg ml^{-1} of proteinase K and analysed by non-denaturing gel electrophoresis on 10% polyacrylamide gels. A free biotin (Sigma-Aldrich) trap was also included with the added helicase to give a final biotin concentration of 100 µM[59]. Gels were quantified as previously published[33].

Ensemble Steady-State and Time-Resolved Fluorescence Spectroscopy. To verify the donor: acceptor ratio and protein concentration, absorption spectra were acquired using a Cary 60 UV-Vis spectrophotometer (Agilent Technologies) over a 200-800 nm wavelength range with 0.5 nm intervals. Rep concentrations were estimated Beer-Lambert's law using extinction coefficients 112,000 $M^{-1}cm^{-1}$ and 270,000 $M^{-1}cm^{-1}$ for Alexa546 and Alexa647, respectively. Steady-state and time-resolved fluorescence spectroscopy was performed using a FluoTime300 spectrophotometer equipped with a hybrid PMT detector (PMA Hybrid 07, Picoquant). Fluorescence emission spectra were measured using continuous 532 nm wavelength laser excitation under magic angle conditions. Ensemble FRET efficiencies, E, were estimated as E = A /(D+A), where D and A represent integrated spectral intensities between 550-650 nm and 650-800 nm, corresponding to donor and acceptor emission, respectively. Time-resolved fluorescence decays were recorded by time-correlated single-photon counting (TCSPC) using excitation from a picosecond pulsed laser (Picoquant, LDH-D-FA-530L) coupled to the entry port of FluoTime300 spectrophotometer. Samples were excited with 30 MHz repetition rate. Fluorescence emission decays at 572 nm were collected until 10^3 counts accumulated at the decay maximum. Fluorescence decay curves were fitted by iterative re-convolution of the instrument response function (IRF) and the observed fluorescence decay assuming a multi-exponential decay function:

$$I_t = \sum_{i=1}^{n} \alpha_{ie^{-t/\tau_i}}$$

I_t is the fluorescence intensity as a function of time, t, (normalized to the intensity at t=0); τ_i the fluorescence lifetime of the ith decay component, α_i is the fractional amplitude of that component. The fit quality was judged based on convergence of the reduced Chi squared. To determine anisotropy and rotational correlation times, the following model was applied to time-resolved decays obtained under polarization conditions:

$$y = \int_{-\infty}^{t} [dt'[IRF(t - Shift_{IRF}) - Bkgr_{IRF}]I(\theta, t - t')] + Bkgr_{dec}$$

$$I(\theta, t) = \left[\frac{k(\theta)}{3}\right] \sum_{i=1}^{3} A_i e^{-\frac{t}{t_i}} \left[1 + (3\cos^2\theta - 1)(re^{-\frac{t}{\varphi}} + R_{inf}\right]$$

A_i is the amplitude of the ith component, t_i is the lifetime of the ith component, r is the initial anisotropy, φ is the correlation time, $k(\vartheta)$ is the experimental matching factor for polarized fluorescence decay collected at emission polarizer angle ϑ, $Bkgr_{dec}$ is the time-independent decay background counts, $Bkgr_{IRF}$ is the time-independent instrument response background counts and $Shift_{IRF}$ is the time shift between the IRF and decay.

Fluorescence correlation spectroscopy (FCS). FCS was performed at room temperature in solution of viscosity 0.001 Pa s using a Zeiss LSM 880 microscope equipped with a GaAsP detector. Normalized FCS curves associated with freely-diffusing Alexa546-Alexa647 Rep helicase were recorded using 30 nM solutions in low or high salt buffer (20 mM Tris-HCl pH 7.9) using 514 nm laser excitation wavelength of typical 4 mW power at the sample plane. The confocal volume was measured using a calibration sample of 6 nM Rhodamine-6G at 21°C, diffusion coefficient 400 $\mu m^2 s^{-1}$ [77]. For qualitative comparison, correlation curves G(τ) could be fitted to heuristic diffusion models incorporated in the instrumentation software, with buffer solutions assumed purely viscous as opposed to viscoelastic[78],

and the variation in solution viscosity due to the range of ionic strength used was assumed to be negligible[79].

VAEM and TIRF microscopy. Our strategy for solution-based imaging of single-molecule doubly labelled surface-immobilized Rep was based on site-specific and ionic interaction between His-tagged Rep and an anti penta-His tag antibody coupled to a glass coverslip via biotin-avidin[80]. Biotinylated polyethylene glycol (PEG) non-specifically adsorbed to pre-silanized surfaces binds to the multivalent avidin protein, providing a specific binding site for the biotinylated anti-His tag which ionically attaches to 6X His antibody on Rep's N-terminus, allowing for conformational freedom for Rep. We used a home-built system enabling control of the angle of incidence by lateral displacement of a downstream lens on a micrometer lateral translation mount[81] [81,82]to generate objective lens mediated VAEM/TIRF with molecular detection precision. The angle of incidence was typically set at approximately 45˚ for VAEM, with depth of penetration of approximately 100 nm for TIRF, excitation intensity at the sample set at 160 W cm^{-2}. Lasers used were Obis LS 532 nm and Obis LX 640 nm at nominal 5 mW power. Fluorescence was collected through a Nikon CFI Apochromat TIRF 100XC oil objective lens, dichroic mirror (Chroma, ZT532/640rpc), propagated through a Photometrics Dual-View colour splitter (dichroic mirror cut-off wavelength 640 nm, donor channel emission bandpass filter of 585/65 nm, and acceptor channel 647 nm long pass) and imaged at an exposure time of 50 ms per frame onto a Prime 95B sCMOS camera at magnification 50nm/pixel. Recorded images were analysed in MATLAB (R2019a) using iSMS single-molecule FRET microscopy software[83] . Briefly, emission channels were aligned, and background-corrected donor and acceptor fluorescence trajectories obtained by integration of pixel intensity within a 9 x 9-pixel area centre on the detected fluorescent foci signals for each time point. Apparent FRET efficiencies were then calculated as described previously.

Confocal Microscopy smFRET. Freely diffusing smFRET microscopy was performed using a home built confocal microscope as previously described[84]. Individual photons were time-stamped to 10 ns precision. The ALEX cycle period was set nominally to 100 µs (comprising 45 µs donor excitation, 5 µs lasers off, 45 µs acceptor excitation, 5 µs lasers off). Individual photon bursts were typically ~1 ms long. Labelled Rep was diluted to ~50 pM in either low of high salt buffer and imaged in 1-hour acquisitions. Multiple hour-long acquisitions were collated (7 and 6 for low and high respectively) to collect >1,000 doubly labelled bursts for each condition. Data were analysed using FRETBursts[85] in Jupyter notebooks. After background determination and subtraction, doubly labelled bursts were identified using a dual channel burst search (minimum 50 photons in each channel) and accurate FRET correction factors were determined[86]. Fully corrected FRET efficiencies were then fit to a 4 Gaussian model.

BVA and H2MM analysis of smFRET confocal microscopy data. To investigate if freely diffusing Rep was exhibiting millisecond timescale dynamics, the BVA[62] was used using MATLAB program PAM[87]. For H2MM analysis, first an all photons burst search (APBS) was performed with the rate threshold for bursts (F) set at 6 times greater than the background count rate, AA photons ≥50, burst size ≥50, and stoichiometry between 0.2 and 0.8. Using the H2MM_C package[88] models were optimised for 1 to *n*+1 states, where *n* is the lowest number of states to achieve the BIC' < 0.005 threshold. The set ALEX cycle period of 100 µs imposed a hypothetical upper limit for rate constant estimates of 1,000 s^{-1}, however, the transition analysis was capable of pushing beyond this by looking at photon arrival times directly within a burst.

ABEL trapping, sample preparation and analysis. The ABEL trap design was similar to previous reports[60,64,89]. 532 nm wavelength laser excitation was provided by a pulsed (60 MHz, ~100 ps pulse width) supercontinuum laser (Leukos ROCK 400-4) paired with a computer-controllable acousto-optic tunable filter (AOTF; Leukos TANGO VIS). A pair of acousto-optic deflectors (AODs) (MTT110-B50A1.5-VIS) arranged orthogonally and controlled by a direct digital synthesizer (DDSPA2X-D8b15b-34) were used to scan the focused laser spot at the sample in a ~3x3 μm^2 optimized knight's tour pattern[90]. The

sample region was aligned confocal to two avalanche photodiodes (Excelitas SPCM-AQRH-14-ND) in two color channels. Emission from each trapped molecule passes through the dichroic in the microscope (Chroma ZT532/640rpc-UF1), gets filtered through a 300 µm diameter pinhole, is split via a dichroic mirror (Chroma T610lpxr) into green and red channels which are further split into parallel and perpendicular emission channels using polarizing beam splitters (Thorlabs PBS251). Filters in each detection channel are as follows: Green Channel: Semrock BLP01-532R-25, Chroma ET542lp, FF01-571/71-25; Red Channel: Chroma ET655lp, Semrock FF01-660/52-25.

Upon detecting a photon, the APDs send a signal to the TCSPC (Picoquant Multiharp 150) and to an FPGA (NI PCIe-78656) for closed-loop feedback. For each photon, a position estimate was determined based on the laser position when the photon was detected and a pre-calibrated lag. A Kalman filter was used to reduce noise in position estimates [90]. X and Y feedback voltages, proportional to distance from the centre, were applied via the FPGA output to a 10x voltage amplifier (Pendulum F10AD) attached to four platinum electrodes placed orthogonally in X and Y reservoirs of the microfluidic cell (in-house fabrication similar to Cohen and Moerner (2005)[91]). Immediately before trapping, labeled protein was diluted in a buffer containing 10 mM tricine (pH 8.0), 1.25 mM $MgCl_2$, 30 µM KCl, 100 µg/ml BSA, glycerol at 5% for high salt experiments and 10% for low salt experiments, an oxygen scavenging system (60 nM protocatechuate-3,4-dioxygenase and 2.6 mM protocatechuic acid) to final concentration of approximately 700 pM, and NaCl added (25 mM for low salt; 500 mM for high salt). To prevent protein from sticking to the microfluidic cells, cells were passivated using mPEG-silane (Laysan Bio MPEG-SIL-5000, 50 mg ml^{-1}) in 95% ethanol, 5% water, and 10 mM acetic acid[92], incubated for approximately 48 hours then in 1% Tween for 10 minutes and rinsed with ultrapure water.

Data analysis was performed using custom software written in MATLAB (Mathworks). Photon arrival times recorded by the Multiharp were used to construct binned time traces. We subtracted the background counts from each of the emission channels. Background bins were identified using an information-criteria-optimized (AIC) K-means clustering algorithm. Bins of constant brightness were determined using a changepoint algorithm that identifies the bins where a substantial change in brightness occurs[93]. Bins between two changepoints are considered a level and get averaged to assign a single brightness or E_{FRET} value for the whole level. The following corrections were performed to the calculated E_{FRET}: Donor leakage into the acceptor channel was removed using the emission detected in the acceptor channel for donor only molecules. The differences in detection efficiency of the parallel and perpendicular channels were corrected using the *G* factor. We also corrected for differences in transmission of donor and acceptor emission based on the filters used (71 % transmission for AF546; 55 % transmission for AF647). Lifetime delays for all photons in each identified level were collected into a histogram and fit with a single exponential decay lifetime convolved with the measured instrument response function, using a maximum likelihood algorithm after Zander et al.[94] and Brand et al.[95] as previously described[61,63,64]. Calculated parameters from levels longer than 30 ms are plotted as scatter heatmaps or histograms as shown in Fig. 3.

MD Simulations. Simulations were generated using AMBER 18 [96] with CUDA supported MD engine, pmemd, using the ff14SB and bsc1 forcefields for the protein and DNA parameters, respectively[97,98] Open and closed Rep starting structures were from PDB ID 1UAA[24]. Each structure had missing amino acids in the loop connecting 2A and 2B subdomains, with the closed structure missing eight (M539, M540, E541, R542, G543, E544, S545, E546) and three in the open (G543, E544, S545). These gaps were filled by incorporating the loop from the closed structure into the open, with the remaining three amino acid gaps in both structures filled by loops from the homologous protein UvRD (PDB ID 2IS2). The protonated state was determined using the Virginia Tech H++ web server[99]. Simulations were carried out on open and closed structures separately with and without DNA at 300 K and 1 atm[100]. The system was solvated implicitly using the generalised Born model with GBneck2 corrections with an infinite cut-off, set with mbondi3 radii. There were two monovalent salt conditions, low (0.01 M) and high (0.5 M). Simulations were minimised and equilibrated using standard protocols[101] , run for 50 ns.

Temperature was maintained at 300 K with a Langevin thermostat using collision frequency 1.0 ps^{-1}. Four replicas were run for each of the eight conditions: the protein with/without DNA starting from open/closed in high/low salt, totalling 32 simulations and 1.6 µs. The Cpptraj program was used to calculate all analysis including RMSD, radius of gyration and frame clustering[102]. The latter was performed via hierarchical agglomerative clustering based on the average linkage algorithm using RMSd between frames as a distance metric. Angles between 2B and 2A were measured using the centre of mass of each subdomain.

FRET accessible volume (AV) analysis

AV and FRET efficiency predictions from the MD simulations were calculated using the FRET modelling software Olga[68]. Fluorophores with their respective linkers were implemented onto modelling structures of Rep with the following parameters:

Parameter	Alexa Fluor 546 (donor)	Alexa Fluor 647 (acceptor)
Linker length (Å)	14	14
Linker width (Å)	4.5	4.5
Radius 1 (Å)	6.8	11
Radius 2 (Å)	3.9	3
Radius 3 (Å)	1.8	1.5
R_0 (Å)	68.2	

Data availability

Data used in this article can be accessed from https://doi.org/10.5281/zenodo.12666478.

References

1. Berezhna, S. Y., Gill, J. P., Lamichhane, R. & Millar, D. P. Single-molecule Förster resonance energy transfer reveals an innate fidelity checkpoint in DNA polymerase i. *J Am Chem Soc* **134**, 11261–11268 (2012).
2. Santoso, Y. *et al.* Conformational transitions in DNA polymerase I revealed by single-molecule FRET. *Proc Natl Acad Sci U S A* **107**, 715–720 (2010).
3. Hohlbein, J. *et al.* Conformational landscapes of DNA polymerase I and mutator derivatives establish fidelity checkpoints for nucleotide insertion. *Nature Communications 2013 4:1* **4**, 1–11 (2013).
4. Evans, G. W., Hohlbein, J., Craggs, T., Aigrain, L. & Kapanidis, A. N. Real-time single-molecule studies of the motions of DNA polymerase fingers illuminate DNA synthesis mechanisms. *Nucleic Acids Res* **43**, 5998–6008 (2015).
5. Singleton, M. R., Dillingham, M. S. & Wigley, D. B. Structure and mechanism of helicases and nucleic acid translocases. *Annu. Rev. Biochem.* **76**, 23–50 (2007).
6. van Eijk, E., Wittekoek, B., Kuijper, E. J. & Smits, W. K. DNA replication proteins as potential targets for antimicrobials in drug-resistant bacterial pathogens. *Journal of Antimicrobial Chemotherapy* **72**, 1275 (2017).

7. Painter, R. E. *et al.* Elucidation of DnaE as the Antibacterial Target of the Natural Product, Nargenicin. *Chem Biol* **22**, 1362–1373 (2015).
8. Shadrick, W. R. *et al.* Discovering New Medicines Targeting Helicases: Challenges and Recent Progress. *Slas Discovery* **18**, 761 (2013).
9. von Hippel, P. H. & Delagoutte, E. A general model for nucleic acid helicases and their 'coupling' within macromolecular machines. *Cell* **104**, 177–190 (2001).
10. Mackintosh, S. G. & Raney, K. D. DNA unwinding and protein displacement by superfamily 1 and superfamily 2 helicases. *Nucleic Acids Res.* **34**, 4154–4159 (2006).
11. Finkelstein, I. J., Visnapuu, M. L. & Greene, E. C. Single-molecule imaging reveals mechanisms of protein disruption by a DNA translocase. *Nature* **468**, 983–987 (2010).
12. Terakawa, T., Redding, S., Silverstein, T. D. & Greene, E. C. Sequential eviction of crowded nucleoprotein complexes by the exonuclease RecBCD molecular motor. *Proc. Natl. Acad. Sci. U S A* **114**, E6322–E6331 (2017).
13. Morris, P. D. & Raney, K. D. DNA helicases displace streptavidin from biotin-labeled oligonucleotides. *Biochemistry* **38**, 5164–51671 (1999).
14. Sokoloski, J. E., Kozlov, A. G., Galletto, R. & Lohman, T. M. Chemo-mechanical pushing of proteins along single-stranded DNA. *Proc. Natl. Acad. Sci. U S A* **113**, 6194–6199 (2016).
15. Ivessa, A. S. *et al.* The Saccharomyces cerevisiae helicase Rrm3p facilitates replication past nonhistone protein-DNA complexes. *Mol. Cell* **12**, 1525–1536 (2003).
16. Guy, C. P. *et al.* Rep Provides a Second Motor at the Replisome to Promote Duplication of Protein-Bound DNA. *Mol. Cell* **36**, 654–666 (2009).
17. Boubakri, H., de Septenville, A. L., Viguera, E. & Michel, B. The helicases DinG, Rep and UvrD cooperate to promote replication across transcription units in vivo. *EMBO J.* **29**, (2010).
18. Killelea, T., Hawkins, M., Howard, J. L. J. L. J. L., McGlynn, P. & Bolt, E. L. E. L. DNA replication roadblocks caused by Cascade interference complexes are alleviated by RecG DNA repair helicase. *RNA Biol* **16**, 543–548 (2019).
19. Hawkins, M. *et al.* Direct removal of RNA polymerase barriers to replication by accessory replicative helicases. *Nucleic Acids Res* **47**, (2019).
20. Paul, T. *et al.* E. coli Rep helicase and RecA recombinase unwind G4 DNA and are important for resistance to G4-stabilizing ligands. *Nucleic Acids Res* **48**, 6640–6653 (2020).
21. Lestini, R. & Michel, B. UvrD and UvrD252 counteract RecQ, RecJ, and RecFOR in a rep mutant of Escherichia coli. *J. Bacteriol.* **190**, 5995–6001 (2008).
22. Dillingham, M. S., Soultanas, P., Wiley, P., Webb, M. R. & Wigley, D. B. Defining the roles of individual residues in the single-stranded DNA binding site of PcrA helicase. *Proceedings of the National Academy of Sciences* **98**, 8381–8387 (2001).
23. Myong, S., Rasnik, I., Joo, C., Lohman, T. M. & Ha, T. Repetitive shuttling of a motor protein on DNA. *Nature* **437**, 1321–1325 (2005).
24. Korolev, S., Hsieh, J., Gauss, G. H., Lohman, T. M. & Waksman, G. Major domain swiveling revealed by the crystal structures of complexes of E. coli Rep helicase bound to single-stranded DNA and ADP. *Cell* **90**, 635–647 (1997).
25. Jia, H. *et al.* Rotations of the 2B Sub-domain of E. coli UvrD Helicase/Translocase Coupled to Nucleotide and DNA Binding. *J. Mol. Biol.* **411**, 633–648 (2011).
26. Arslan, S., Khafizov, R., Thomas, C. D., Chemla, Y. R. & Ha, T. Engineering of a superhelicase through conformational control. *Science (1979)* **348**, 344–347 (2015).

27. Comstock, M. J. *et al.* Direct observation of structure-function relationship in a nucleic acid-processing enzyme. *Science (1979)* **348**, 352–354 (2015).
28. Nguyen, B., Ordabayev, Y., Sokoloski, J. E., Weiland, E. & Lohman, T. M. Large domain movements upon UvrD dimerization and helicase activation. *Proceedings of the National Academy of Sciences* **114**, 12178–12183 (2017).
29. Atkinson, J., Gupta, M. K. & McGlynn, P. Interaction of Rep and DnaB on DNA. *Nucleic Acids Res* **39**, 1351–1359 (2011).
30. Atkinson, J. *et al.* Localization of an accessory helicase at the replisome is critical in sustaining efficient genome duplication. *Nucleic Acids Res* **39**, 949–957 (2011).
31. Gupta, M. K. *et al.* Protein-DNA complexes are the primary sources of replication fork pausing in Escherichia coli. *Proceedings of the National Academy of Sciences* **110**, 7252–7257 (2013).
32. Syeda, A. H. *et al.* Single-molecule live cell imaging of Rep reveals the dynamic interplay between an accessory replicative helicase and the replisome. *Nucleic Acids Res* **47**, 6287–6298 (2019).
33. Brendza, K. M. *et al.* Autoinhibition of Escherichia coli Rep monomer helicase activity by its 2B subdomain. *Proc. Natl. Acad. Sci. U S A* **102**, 10076–10081 (2005).
34. Cheng, W., Hsieh, J., Brendza, K. M. & Lohman, T. M. E. coli Rep Oligomers are Required to Initiate DNA Unwinding in vitro. *J. Mol. Biol.* **310**, 327-350. (2001).
35. Cheng, W. *et al.* The 2B domain of the Escherichia coli Rep protein is not required for DNA helicase activity. *Proc. Natl. Acad. Sci. U S A* **99**, 16006–16011 (2002).
36. Brüning, J. G., Howard, J. A. L., Myka, K. K., Dillingham, M. S. & McGlynn, P. The 2B subdomain of Rep helicase links translocation along DNA with protein displacement. *Nucleic Acids Res* **46**, 8917–8925 (2018).
37. Velankar, S. S., Soultanas, P., Dillingham, M. S., Subramanya, H. S. & Wigley, D. B. Crystal structures of complexes of PcrA DNA helicase with a DNA substrate indicate an inchworm mechanism. *Cell* **97**, 75–84 (1999).
38. Lee, J. Y. & Yang, W. UvrD helicase unwinds DNA one base pair at a time by a two-part power stroke. *Cell* **127**, 1349–1360 (2006).
39. Ha, T. *et al.* Initiation and re-initiation of DNA unwinding by the Escherichia coli Rep helicase. *Nature* **419**, 638–641 (2002).
40. Maluf, N. K., Fischer, C. J. & Lohman, T. M. A Dimer of Escherichia coli UvrD is the active form of the helicase in vitro. *J. Mol. Biol.* **325**, 913–935 (2003).
41. Niedziela-Majka, A., Chesnik, M. A., Tomko, E. J. & Lohman, T. M. Bacillus stearothermophilus PcrA monomer is a single-stranded DNA translocase but not a processive helicase in vitro. *J. Biol. Chem.* **282**, 27076–27085 (2007).
42. Lee, K. S., Balci, H., Jia, H., Lohman, T. M. & Ha, T. Direct imaging of single UvrD helicase dynamics on long single-stranded DNA. *Nat. Commun.* **4**, 1878 (2013).
43. Ordabayev, Y. A., Nguyen, B., Kozlov, A. G., Jia, H. & Lohman, T. M. UvrD helicase activation by MutL involves rotation of its 2B subdomain. *Proc Natl Acad Sci U S A* **116**, 16320–16325 (2019).
44. Ordabayev, Y. A., Nguyen, B., Niedziela-Majka, A. & Lohman, T. M. Regulation of UvrD Helicase Activity by MutL. *J Mol Biol* **430**, 4260–4274 (2018).
45. Chisty, L. T. *et al.* Monomeric PcrA helicase processively unwinds plasmid lengths of DNA in the presence of the initiator protein RepD. *Nucleic Acids Res.* **41**, 5010–5023 (2013).

46. Dienerowitz, M., Howard, J. A. L., Quinn, S. D., Dienerowitz, F. & Leake, M. C. Single-molecule FRET dynamics of molecular motors in an ABEL trap. *Methods* **193**, 96–106 (2021).
47. Sisamakis, E., Valeri, A., Kalinin, S., Rothwell, P. J. & Seidel, C. A. M. Accurate single-molecule FRET studies using multiparameter fluorescence detection. *Methods Enzymol* **475**, 455–514 (2010).
48. Yu, L. *et al.* A Comprehensive Review of Fluorescence Correlation Spectroscopy. *Front Phys* **9**, 644450 (2021).
49. Bacia, K., Kim, S. A. & Schwille, P. Fluorescence cross-correlation spectroscopy in living cells. *Nat Methods* **3**, 83–89 (2006).
50. Miller, H. *et al.* High-Speed Single-Molecule Tracking of CXCL13 in the B-Follicle. *Front Immunol* **9**, 1073 (2018).
51. Xi, X. G. & Deprez, E. Monitoring helicase-catalyzed DNA unwinding by fluorescence anisotropy and fluorescence cross-correlation spectroscopy. *Methods* **51**, 289–294 (2010).
52. Li, N. *et al.* Multiple Escherichia coli RecQ helicase monomers cooperate to unwind long DNA substrates: A fluorescence cross-correlation spectroscopy study. *Journal of Biological Chemistry* **285**, 6922–6936 (2010).
53. Leake, M. C. *Biophysics : Tools and Techniques for the Physics of Life*. (CRC Press, 2024).
54. Shashkova, S. & Leake, M. C. Single-molecule fluorescence microscopy review: shedding new light on old problems. *Biosci Rep* **37**, (2017).
55. Sasmal, D. K., Pulido, L. E., Kasal, S. & Huang, J. Single-molecule fluorescence resonance energy transfer in molecular biology. *Nanoscale* **8**, 19928–19944 (2016).
56. Leake, M. C. The physics of life: one molecule at a time. *Philos Trans R Soc Lond B Biol Sci* **368**, 20120248 (2013).
57. Miller, H., Zhou, Z., Shepherd, J., Wollman, A. J. M. M. & Leake, M. C. Single-molecule techniques in biophysics: a review of the progress in methods and applications. *Reports on Progress in Physics* **81**, 024601 (2017).
58. Konopka, C. A. & Bednarek, S. Y. Variable-angle epifluorescence microscopy: a new way to look at protein dynamics in the plant cell cortex. *Plant J* **53**, 186–196 (2008).
59. Cohen, A. E. & Moerner, W. E. Suppressing Brownian motion of individual biomolecules in solution. *Proc Natl Acad Sci U S A* **103**, 4362–4365 (2006).
60. Squires, A. H. *et al.* Single-molecule trapping and spectroscopy reveals photophysical heterogeneity of phycobilisomes quenched by Orange Carotenoid Protein. *Nat Commun* **10**, (2019).
61. Goldsmith, R. H. & Moerner, W. E. Watching conformational- and photo-dynamics of single fluorescent proteins in solution. *Nat Chem* **2**, 179–186 (2010).
62. Wang, Q. & Moerner, W. E. An Adaptive Anti-Brownian ELectrokinetic trap with real-time information on single-molecule diffusivity and mobility. *ACS Nano* **5**, 5792–5799 (2011).
63. Wang, Q. & Moerner, W. E. Lifetime and spectrally resolved characterization of the photodynamics of single fluorophores in solution using the anti-Brownian electrokinetic trap. *J Phys Chem B* **117**, 4641–4648 (2013).
64. Squires, A. H. & Moerner, W. E. Direct single-molecule measurements of phycocyanobilin photophysics in monomeric C-phycocyanin. *Proc Natl Acad Sci U S A* **114**, 9779–9784 (2017).

65. Torella, J. P., Holden, S. J., Santoso, Y., Hohlbein, J. & Kapanidis, A. N. Identifying Molecular Dynamics in Single-Molecule FRET Experiments with Burst Variance Analysis. *Biophys J* **100**, 1568–1577 (2011).
66. Pirchi, M. *et al.* Photon-by-Photon Hidden Markov Model Analysis for Microsecond Single-Molecule FRET Kinetics. *Journal of Physical Chemistry B* **120**, 13065–13075 (2016).
67. Lerner, E., Ingargiola, A. & Weiss, S. Characterizing highly dynamic conformational states: The transcription bubble in RNAP-promoter open complex as an example. *Journal of Chemical Physics* **148**, 123315 (2018).
68. Dimura, M. *et al.* Automated and optimally FRET-assisted structural modeling. *Nature Communications 2020 11:1* **11**, 1–14 (2020).
69. Fixman, M. Radius of gyration of polymer chains. II. Segment density and excluded volume effects. *J Chem Phys* **36**, 3123–3129 (1962).
70. Brüning, J. G., Howard, J. L. & McGlynn, P. Accessory replicative helicases and the replication of protein-bound DNA. *J Mol Biol* **426**, 3917–3928 (2014).
71. Makurath, M. A., Whitley, K. D., Nguyen, B., Lohman, T. M. & Chemla, Y. R. Regulation of Rep helicase unwinding by an auto-inhibitory subdomain. *Nucleic Acids Res* **47**, 2523–2532 (2019).
72. Wong, I., Chao, K. L., Bujalowski, W. & Lohman, T. M. DNA-induced dimerization of the Escherichia coli rep helicase. Allosteric effects of single-stranded and duplex DNA. *J. Biol. Chem.* **267**, 7596–7610 (1992).
73. Wong, I. & Lohman, T. M. A two-site mechanism for ATP hydrolysis by the asymmetric Rep dimer P2S as revealed by site-specific inhibition with ADP-A1F4. *Biochemistry* **36**, 3115–3125 (1997).
74. Chao, K. L. & Lohman, T. M. DNA-induced dimerization of the Escherichia coli Rep helicase. *J. Mol. Biol.* **221**, 1165–1181 (1991).
75. Nguyen, B., Shinn, M. K., Weiland, E. & Lohman, T. M. Regulation of E. coli Rep helicase activity by PriC. *J Mol Biol* **433**, 167072 (2021).
76. Rasnik, I., Myong, S., Cheng, W., Lohman, T. M. & Ha, T. DNA-binding Orientation and Domain Conformation of the E. coli Rep Helicase Monomer Bound to a Partial Duplex Junction: Single-molecule Studies of Fluorescently Labeled Enzymes. *J Mol Biol* **336**, 395–408 (2004).
77. Gendron, P. O., Avaltroni, F. & Wilkinson, K. J. Diffusion coefficients of several rhodamine derivatives as determined by pulsed field gradient-nuclear magnetic resonance and fluorescence correlation spectroscopy. *J Fluoresc* **18**, 1093–1101 (2008).
78. Linke, W. A. & Leake, M. C. Multiple sources of passive stress relaxation in muscle fibres. *Phys Med Biol* **49**, 3613–27 (2004).
79. Zhang, H. L. & Han, S. J. Viscosity and density of water+sodium chloride+potassium chloride solutions at 298.15 K. *J Chem Eng Data* **41**, 516–520 (1996).
80. Pal, P., Lesoine, J. F., Lieb, M. A., Novotny, L. & Knauf, P. A. A novel immobilization method for single protein spFRET studies. *Biophys J* **89**, L11–L13 (2005).
81. Llorente-Garcia, I. *et al.* Single-molecule in vivo imaging of bacterial respiratory complexes indicates delocalized oxidative phosphorylation. *Biochim Biophys Acta* **1837**, 811–24 (2014).
82. Dresser, L. *et al.* Amyloid-β oligomerization monitored by single-molecule stepwise photobleaching. *Methods* **193**, 80–95 (2021).

83. Preus, S., Noer, S. L., Hildebrandt, L. L., Gudnason, D. & Birkedal, V. iSMS: single-molecule FRET microscopy software. *Nat Methods* **12**, 593–594 (2015).
84. Ambrose, B. *et al.* The smfBox is an open-source platform for single-molecule FRET. *Nature Communications 2020 11:1* **11**, 1–6 (2020).
85. Ingargiola, A., Lerner, E., Chung, S. Y., Weiss, S. & Michalet, X. FRETBursts: An Open Source Toolkit for Analysis of Freely-Diffusing Single-Molecule FRET. *PLoS One* **11**, e0160716 (2016).
86. Hellenkamp, B. *et al.* Precision and accuracy of single-molecule FRET measurements—a multi-laboratory benchmark study. *Nat Methods* **15**, 669–676 (2018).
87. Schrimpf, W., Barth, A., Hendrix, J. & Lamb, D. C. PAM: A Framework for Integrated Analysis of Imaging, Single-Molecule, and Ensemble Fluorescence Data. *Biophys J* **114**, 1518–1528 (2018).
88. Harris, P. D. *et al.* Multi-parameter photon-by-photon hidden Markov modeling. *Nature Communications 2022 13:1* **13**, 1–12 (2022).
89. Wang, Q. & Moerner, W. E. Dissecting pigment architecture of individual photosynthetic antenna complexes in solution. *Proc Natl Acad Sci U S A* **112**, 13880–13885 (2015).
90. Wang, Q. & Moerner, W. E. Optimal strategy for trapping single fluorescent molecules in solution using the ABEL trap. *Appl Phys B* **99**, 23–30 (2010).
91. Cohen, A. E. & Moerner, W. E. An all-glass microfluidic cell for the ABEL trap: fabrication and modeling. *https://doi.org/10.1117/12.622418* **5930**, 191–198 (2005).
92. Wilson, H. & Wang, Q. ABEL-FRET: tether-free single-molecule FRET with hydrodynamic profiling. *Nat Methods* **18**, 816–820 (2021).
93. Watkins, L. P. & Yang, H. Detection of intensity change points in time-resolved single-molecule measurements. *J Phys Chem B* **109**, 617–628 (2005).
94. Zander, C. *et al.* Detection and characterization of single molecules in aqueous solution. *Appl Phys B* **63**, 517–523 (1996).
95. Brand, L., Eggeling, C., Zander, C., Drexhage, K. H. & Seidel, C. A. M. Single-molecule identification of Coumarin-120 by time-resolved fluorescence detection: Comparison of one- and two-photon excitation in solution. *Journal of Physical Chemistry A* **101**, 4313–4321 (1997).
96. D.A. Case, R.M. Betz, D.S. Cerutti, T.E. Cheatham, III, T.A. Darden, R.E. Duke, T.J. Giese, H. Gohlke, A.W. Goetz, N. Homeyer, S. Izadi, P. Janowski, J. Kaus, A. Kovalenko, T.S. Lee, S. LeGrand, P. Li, C.Lin, T. Luchko, R. Luo, B. Madej, D. Mermelstein, K, S. F. Amber 2016. Preprint at (2016).
97. Maier, J. A. *et al.* ff14SB: Improving the Accuracy of Protein Side Chain and Backbone Parameters from ff99SB. *J Chem Theory Comput* **11**, 3696–3713 (2015).
98. Ivani, I. *et al.* Parmbsc1: a refined force field for DNA simulations. *Nature Methods 2016 13:1* **13**, 55–58 (2015).
99. Anandakrishnan, R., Aguilar, B. & Onufriev, A. V. H++ 3.0: Automating pK prediction and the preparation of biomolecular structures for atomistic molecular modeling and simulations. *Nucleic Acids Res* **40**, 537–541 (2012).
100. Berendsen, H. J. C., Postma, J. P. M., Van Gunsteren, W. F., Dinola, A. & Haak, J. R. Molecular dynamics with coupling to an external bath. *J Chem Phys* **81**, 3684–3690 (1984).

101. Watson, G. D., Chan, E. W., Leake, M. C. & Noy, A. Structural interplay between DNA-shape protein recognition and supercoiling: The case of IHF. *Comput Struct Biotechnol J* **20**, 5264–5274 (2022).
102. Roe, D. R. & Cheatham, T. E. PTRAJ and CPPTRAJ: Software for processing and analysis of molecular dynamics trajectory data. *J Chem Theory Comput* **9**, 3084–3095 (2013).

Acknowledgements

We thank the Bioscience Technology Facility, University of York for technical support. MCL supported by BBSRC (grant BB/P000746/1). AHS acknowledges research support from the Neubauer Family Foundation and NSF QLCI QuBBE grant OMA-2121044. AN acknowledges EPSRC (EP/N027639/1) and University of York Research Priming funds [ref 50109444 and Center for Future Health 204829]. SQ acknowledges support from Alzheimer's Research UK (RF2019-A-001). AA-B was supported by EPSRC grant EP/N509802/1. The authors acknowledge EP/T022205/1, EP/R029407/1, EP/P020259/1 and EP/V034030/1 and the local York facilities for computational resources. MD acknowledges financial support from the DFG - German Research Foundation (528591139- FIP 31) and the Carl Zeiss Stiftung (P2022-13-017). TDC was supported by BBSRC (BB/W00061X/1, BB/T008032/1) and EPSRC (EP/V034804/1).

Author contributions

M.L., J.A.L.H. & S.D.Q. conceived the study. J.A.L.H., S.D.Q., L.D., B.A., M.A.S.A., M.D., A.E. & A.S. were responsible for data collection and analysis. A.N. conceptualised the molecular dynamics studies, which were carried out and analysed by A.A-B. The initial bulk of the manuscript was written by J.A.L.H, S.D.Q. & M.L with contribution, discussion and revision from all of the authors. M.L. was the project administrator.

Competing interests

All authors declare that they have no competing interests.