








# Red $\beta$ <sub>177</sub> annealase structure reveals details of oligomerization and $\lambda$ Red-mediated homologous DNA recombination

Received: 4 May 2022

Accepted: 31 August 2022

Published online: 26 September 2022

 Check for updatesTimothy P. Newing <sup>1,2</sup>, Jodi L. Brewster <sup>1,2,3</sup>, Lucy J. Fitschen <sup>1,2,3</sup>, James C. Bouwer <sup>1,3</sup>, Nikolas P. Johnston <sup>1</sup>, Haibo Yu <sup>1,2</sup> & Gökhan Tolun <sup>1,2,3</sup> ✉

The Red $\beta$  protein of the bacteriophage  $\lambda$  red recombination system is a model annealase which catalyzes single-strand annealing homologous DNA recombination. Here we present the structure of a helical oligomeric annealing intermediate of Red $\beta$ , consisting of N-terminal residues 1-177 bound to two complementary 27mer oligonucleotides, determined via cryogenic electron microscopy (cryo-EM) to a final resolution of 3.3 Å. The structure reveals a continuous binding groove which positions and stabilizes complementary DNA strands in a planar orientation to facilitate base pairing via a network of hydrogen bonding. Definition of the inter-subunit interface provides a structural basis for the propensity of Red $\beta$  to oligomerize into functionally significant long helical filaments, a trait shared by most annealases. Our cryo-EM structure and molecular dynamics simulations suggest that residues 133-138 form a flexible loop which modulates access to the binding groove. More than half a century after its discovery, this combination of structural and computational observations has allowed us to propose molecular mechanisms for the actions of the model annealase Red $\beta$ , a defining member of the Red $\beta$ /RecT protein family.

Homologous DNA recombination has long been observed in lambdoid phages. It assists in the formation of chromosomes<sup>1</sup> and is a frequent and common contributor to genetic diversity within phage populations<sup>2,3</sup>. Like many dsDNA viruses, lambdoid phages encode an efficient two-component recombinase system which facilitates recombination between homologous regions of ssDNA through a single-strand annealing (SSA) mechanism<sup>4</sup>. Homologous recombination promoted by viral two-component recombinases has also been shown to increase the efficiency of DNA replication and facilitate the repair of dsDNA breaks throughout the viral life cycle<sup>5,6</sup>.

SSA is an evolutionarily conserved pathway present in many organisms, from simple bacteriophages to humans<sup>7</sup>. However, the first

discovered and best characterized of these systems is the Red recombinase from the *Escherichia coli* bacteriophage lambda (phage  $\lambda$ )<sup>8</sup>. The Red system is comprised of the genes *exo*, *bet* and *gam*, which encode for the proteins  $\lambda$  exonuclease ( $\lambda$ Exo [originally, Red $\alpha$ ]), Red $\beta$  and  $\gamma$ -protein, respectively. The Red system proteins are capable of repairing dsDNA breaks by SSA<sup>9</sup>, while Red $\beta$  alone has been shown to be sufficient for ssDNA recombination<sup>10</sup>.  $\lambda$ Exo is a 25.9 kDa protein that forms a toroidal homotrimer<sup>11</sup>, which binds to dsDNA ends. It resects one of the strands in the 5' to 3' direction, converting the other strand to a 3' ssDNA overhang<sup>12,13</sup>. Red $\beta$  (a.k.a. Beta), a 29.7 kDa protein, binds to and protects this exposed nascent ssDNA, catalyzing its annealing to a homologous DNA molecule<sup>14</sup>.  $\lambda$ Exo and Red $\beta$  have been shown to interact in a 1:1 monomer:monomer stoichiometry to form an

<sup>1</sup>School of Chemistry and Molecular Bioscience, and Molecular Horizons, University of Wollongong, Wollongong, NSW, Australia. <sup>2</sup>Illawarra Health and Medical Research Institute, Wollongong, NSW, Australia. <sup>3</sup>ARC Centre for Cryo-electron Microscopy of Membrane Proteins, University of Wollongong, Wollongong, NSW, Australia. ✉ e-mail: [gokhan\\_tolun@uow.edu.au](mailto:gokhan_tolun@uow.edu.au)

Exonuclease Annealase Two-component Recombinase (EATR) complex (a.k.a. SynExo)<sup>5,15</sup>. The third protein,  $\gamma$  (Gam), though not physically associated with EATR, binds to and inhibits host *E. coli* RecBCD nuclease<sup>16,17</sup> and enhances recombination efficiency<sup>10</sup>.

The phage  $\lambda$  Red system is the archetypal EATR complex and has been studied extensively over the last 50 years. Its high recombination rate has made it a convenient model system to study SSA<sup>9</sup>.  $\lambda$  Red has also emerged as a tool for genetic manipulation, termed recombineering<sup>12,13</sup>. Recombineering has been reported to be significantly more effective than classical bacterial genetic manipulation, with the Red system being key to the development of these technologies<sup>1</sup>. There are currently two proposed models for how SSA may occur, which have been recently reviewed by us<sup>8</sup> and others<sup>5,15</sup>. In brief, variable processing of dsDNA ends by  $\lambda$ Exo leads to the following proposed scenarios: (1) Two partially processed homologous dsDNA molecules are annealed to each other<sup>18</sup>, (2) ssDNA resulting from a partial or complete digestion of dsDNA is annealed to region(s) of transient genomic ssDNA, such as the lagging strand at a DNA replication fork<sup>10,19–21</sup>. In these instances, Red $\beta$  is proposed to mediate homology searching and the annealing of homologous sequences. A version of the latter model forms the basis of another genetic tool, Multiplex Autonomous Genome Engineering (MAGE)<sup>22</sup>, for the simultaneous targeted mutagenesis of numerous genes. However, many questions related to the process of SSA remain unanswered, particularly regarding the detailed molecular mechanisms underlying the catalysis of homology searching and annealing by Red $\beta$ .

Functional homologs of Red $\beta$  appear to be almost ubiquitous, with an increasing number of representatives being discovered across viral, bacterial, and eukaryotic domains. Despite this, few of the annealase proteins have received even rudimentary characterization. Therefore, there is a scarcity of information regarding the biochemical, structural and functional attributes of annealases. However, the available data on Red $\beta$  and other annealases such as ERF (phage P22), ORF6 (Kaposi's Sarcoma Herpes Virus), ICP8 (Herpes Simplex Virus-1), RecT (*E. coli*), human Rad52 and its eukaryotic homologs suggest some commonalities. Annealases are ATP-independent and do not catalyze strand invasion of duplex DNA; thus, they are functionally distinct from the well-known and studied ATPase recombination proteins, such as RecA from *E. coli* and eukaryotic Rad51<sup>9,23,24</sup>. However, in a manner reminiscent of *E. coli* SSB, annealases bind to ssDNA, holding it in an extended conformation, removing secondary structures and protecting it from degradation<sup>25</sup>. The DNA-binding properties of Red $\beta$  differ substantially from those of SSB. Red $\beta$  binds only weakly to ssDNA, but in the presence of magnesium, it binds stably and with high affinity to a nascent dsDNA, the product of annealing, such as one that forms upon the sequential addition of two complementary ssDNA oligos<sup>26</sup>. Red $\beta$  does not bind at all to preformed duplex DNA<sup>26</sup>.

Red $\beta$  is known to form ring-like and filamentous oligomeric superstructures that are proposed to be functionally important. Electron and atomic force microscopy studies have provided images and low-resolution structural information on the oligomeric complexes of Red $\beta$ . In the absence of DNA, full-length Red $\beta$  forms an abundance of small rings or split-lock-washer conformers, each with an average of 11–12 subunits<sup>27,28</sup>. However, in the presence of ssDNA, Red $\beta$  forms a heterogeneous mix of 11–19 subunit rings, distorted and part rings, and short left-handed helical filaments which are seen in both the full length<sup>27,28</sup> and various truncated versions of Red $\beta$ <sup>27–29</sup>. In the presence of heat-denatured complementary DNA, the larger rings and filament superstructures are stabilized and are proposed to represent an annealing intermediate<sup>26–28</sup>. The oligomeric heterogeneity of full-length Red $\beta$  was recently expanded upon through a detailed, native mass spectrometry analysis<sup>30</sup>. Above a protein concentration of 1  $\mu$ M, numerous oligomeric states from 5 to 14 subunits were detected for both Red $\beta$  alone and in the presence of short ssDNA and complementary oligonucleotides. Increasing the Red $\beta$  concentration up to

30  $\mu$ M promoted larger structures, as did an increase in the length of DNA used<sup>30</sup>.

Like Red $\beta$ , other annealase proteins have a propensity to form similar ring-like quaternary structures and helical filament superstructures, which appear to suggest a common mode of action. In addition, rings are often seen at the termini of long helical filaments<sup>27,31</sup>, giving rise to the hypothesis that the ring and helical forms interconvert<sup>27,32</sup>. Many such oligomeric complexes have been observed by electron microscopy, including Orf6<sup>33</sup>, Rad52<sup>24,34–37</sup> and ICP8<sup>31,38–41</sup>. Low-resolution reconstructions of ICP8, both in the form of binary nonameric rings (in the presence of short ssDNA)<sup>32</sup> and bipolar filaments<sup>42</sup> have been generated<sup>32</sup>. The bipolar filaments form in both the absence of DNA<sup>39,41,42</sup> and the presence of heat-denatured DNA<sup>31,40</sup>, and likely represent an intertwining of single filaments seen in the presence of ssDNA<sup>38</sup>. The ring and bipolar filaments have also been proposed as interconverting annealing intermediates<sup>31,32,42</sup>. The fitting of an available crystal structure of the ICP8 monomer, with a 60 residue C-terminal truncation<sup>43</sup>, into the reconstructions has provided information on intermolecular contacts involved in oligomerization.

Though they share functional and oligomerization properties, the annealases of bacterial and phage origin appear to form a group distinct from, and likely evolutionarily unrelated to, the eukaryotic-host viral annealases, such as ICP8, and homologs BALF2 and ORF6. The latter are much larger at ~130 kDa, compared to 29.7 kDa for Red $\beta$ . This difference in size may reflect a diversity of function and a greater number of interacting partners. In addition to involvement in SSA, ICP8 is proposed to have multiple roles in DNA replication, replication compartment formation and gene expression<sup>44,45</sup>.

Phylogenetic analyses of key annealases, including human, bacterial and phage derived proteins, appear to resolve at least three evolutionary lineages, each forming a superfamily: Gp2.5-like, Rad51-like and Rad52-like<sup>46</sup>. Later publications further divided these lineages into five distinct families named for their respective flagship members Sak3, Sak4, Erf, Rad52 and RecT/Red $\beta$ <sup>47</sup>. The only near-atomic resolution structural data of an oligomeric complex available amongst these families is of the undecameric rings of the Rad52 N-terminal domain (residues 1–212)<sup>35,48,49</sup>. These data revealed an inner groove and an outer groove on the ring surface, where ssDNA was bound. The two grooves were suggested to represent two different modes of DNA-binding and reaction intermediates. The deep inner groove, predicted to be the site of annealing, is lined with positively charged arginine and lysine residues that interact with the phosphate backbone of ssDNA, positioning the nucleotides outwards for complementary base pairing.

The architecture of annealase proteins can be crudely separated into two functionally distinct domains. The N-terminal domain (NTD) mediates oligomerization and contains the catalytic functionality. Residues 1–177 of Red $\beta$  (Red $\beta$ <sub>177</sub>) are sufficient for binding to ssDNA<sup>50</sup> and catalyzing the annealing of complementary strands in vitro<sup>29,51</sup>. Although there are some suggestions that the C-terminal domain (CTD) may modulate the strength of DNA interactions, it does not appear to be required<sup>52</sup>. The much smaller CTDs are implicated in mediating protein-protein interactions. Residues 178–261 of Red $\beta$  interact with  $\lambda$ Exo<sup>29</sup> and potentially other host proteins such as SSB<sup>53</sup>. The NTD and CTD of annealases are flexible relative to one another, which presents major challenges for structure determination. The current available crystal structures of both ICP8<sup>43</sup> and Rad52<sup>35,48,49</sup> are C-terminally truncated. However, while these crystal structures provide tantalizing insights into the function of annealase proteins within that family, they are of limited use for broader extrapolation. The amino acid sequence similarity between annealases is very low, often less than 15–20%<sup>46</sup>, making sequence alignments less reliable and impractical for the purpose of generating homology models for annealases in different families.

An X-ray crystal structure of the C-terminal domain of Red $\beta$  (residues 183–261) in complex with the  $\lambda$ Exo toroid, showed a compact 3-helix bundle<sup>15</sup>. However, there is no representative structure of the

NTD, the main domain that carries out the primary functions of DNA binding, annealing and homo-oligomerization required for SSA. As Red $\beta$  is undoubtedly the archetype annealase for its family, the absence of structural information has, until now, been a major barrier in our understanding of not only bacteriophage annealases, but also how SSA functions.

Here we present a structure of the helical annealing intermediate of Red $\beta_{177}$ , in complex with complementary oligonucleotides, determined by cryo-electron microscopy to 3.3 Å. Our structure allows a detailed look at both the nucleoprotein interactions and the inter-protein contacts dictating helical superstructure formation. This structure demonstrates the mechanism of action for Red $\beta_{177}$ , while also having the potential to be used as a template for homology models for other members of this protein superfamily. Combined with CTD crystal structure and the structure prediction we made using AlphaFold v2.0<sup>54</sup>, we finally present a complete picture of the Red $\beta$  structure more than half a century after its discovery.

## Results

We present the structure of the DNA-binding and oligomerization domain of Red $\beta$  annealase from phage  $\lambda$ . This structure, resolved to 3.3 Å, was obtained by cryo-EM via helical reconstruction, using a C-terminally truncated Red $\beta_{177}$  bound to two 27mer complementary oligonucleotides.

### Reconstructed helical cryo-EM density map

The Red $\beta_{177}$  was expressed in *E. coli* and purified by immobilized metal ion affinity chromatography followed by gel filtration chromatography (Supplementary Fig. 1). The purified protein was incubated with oligonucleotides before being plunge frozen for cryo-EM. We collected 4,701 micrographs, from which 259,649 helical particles were identified using template-based helical filament tracing and subsequent particle curation in cryoSPARC<sup>55</sup>. The micrographs contained helices in a variety of conformations, but most were uniform, compact “telephone cord-like” solenoids, though some sections were less condensed (Fig. 1a). The 2D class averages generated during the reconstruction (Fig. 1b and Supplementary Fig. 2a), together with their Fourier transformation (Fig. 1c), clearly show the features of the helical structures formed by Red $\beta_{177}$ . Two subpopulations of particles, representing different helical symmetries, were identified in 2D classification—a 1-start and a 2-start helical assembly, consisting of 26,525 and 77,799 particles, respectively (Supplementary Fig. 1). Previous work has established that Red $\beta$  forms left-handed helices<sup>27</sup>, therefore left-handed helical symmetry was imposed during further processing.

Helical reconstruction and subsequent symmetry determination of the 1-start helical population identified a twist of  $-12.947$  degrees and a rise of 2.078 Å. Following refinement of the local motion and CTF (contrast transfer function), symmetrized, non-uniform helical refinement of the 1-start particles resulted in a map with an estimated final resolution of 3.3 Å (Fig. 1d–f), according to the gold-standard FSC = 0.143 (GSFSC) criterion<sup>56</sup>. However, the local resolution of the map, as determined by cryoSPARC, ranges from 2.8 Å to 4.8 Å (Fig. 1e). The overall image processing workflow can be seen in Supplementary Fig. 3 and structure statistics can be found in Supplementary Table 1. A single turn of the 1-start helix contains 27.8 asymmetric units, each corresponding to a Red $\beta_{177}$  monomer (Fig. 2a). The density of double-stranded DNA, which represents an average of possible DNA sequences bound to a monomer of Red $\beta_{177}$ , can be seen wrapping around the circumference of the protein helix (Figs. 1d, e and 2a). The presence of double-stranded DNA in the map indicates that we have captured an annealing intermediate of the SSA homologous recombination reaction. The attempt to reconstruct the structure of the 2-start helical complex did not yield a reliable map, likely due to the presence of irresolvable continuous variation in the helical symmetry parameters. Therefore, we are presenting only the 2D class averages that indicate a

helical assembly composed of two apparently identical Red $\beta_{177}$  filaments (Supplementary Fig. 2b).

### Atomic model confirms Red $\beta_{177}$ is structurally similar to human Rad52

A total of 161 residues of Red $\beta_{177}$  were built de novo into the density (Fig. 2b). This encompassed residues 3–169, except for a flexible loop region (residues 133–138) which could not be accurately modeled (see below). The density for residues 1–2 and 170–177 was of insufficient quality to allow modeling. However, the quality of the map allowed for the accurate positioning of the side chains of majority of the remaining residues and of the oligonucleotides (Fig. 2c). The coordinates for the atomic model have been deposited into the PDB (7UJL) and the map into EMDB (EMD-26566). In general, the Red $\beta$  N-terminal domain consists of a mixed  $\alpha\beta$  fold consisting of five  $\alpha$ -helices, a  $3_{10}$  helix and five  $\beta$ -strands, which form two separate  $\beta$ -sheet elements (Fig. 3a). The structure contains a recognizable  $\beta\beta\alpha$  core motif ( $\beta$ 3–5 and  $\alpha$ 5), which is a small OB-fold, present in DNA-binding proteins<sup>57</sup>. This motif is also present in the crystal structure of the N-terminal domain of Rad52 (Fig. 3b), where it contributes to the formation of the deep groove of the inner DNA-binding pocket<sup>49</sup>. However, in Red $\beta_{177}$  the  $\beta$ -strands are much shorter, 5–6 amino acids in length, relative to those of the Rad52  $\beta\beta\alpha$  motif (10–14 amino acids in length). As a result, the outward-facing DNA-binding groove on the surface of Red $\beta_{177}$  is much shallower and more exposed than that of Rad52 (Fig. 3a, b). A structural similarity search with Red $\beta_{177}$ , using the DALI server<sup>58</sup>, identified Rad52 as the most similar structural fold. This is consistent with prior work placing Red $\beta$  and many other bacterial annealases in a superfamily that is proposed to share a common fold<sup>46</sup>.

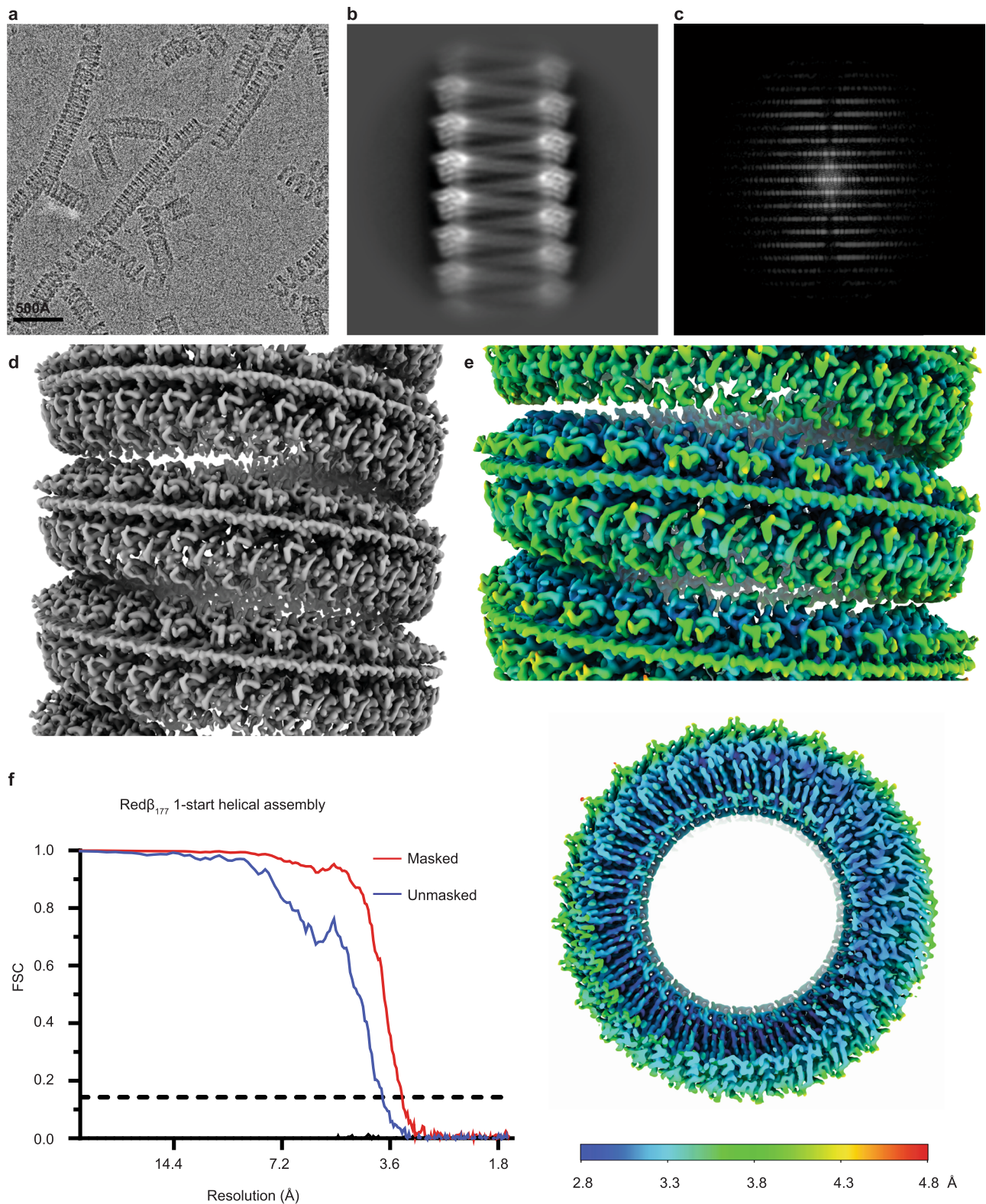
### Multiple amino acid sequence alignment identifies conserved motifs

Using the Red $\beta$  protein sequence (UniProtKB: P03698) as a reference, we constructed a multiple sequence alignment (MSA) based on the top 1000 hits from a BLAST search of the UniProt Reference Cluster (UniRef) 90% database. Following refinement, our MSA consists of 425 reference cluster sequences and is 411 residues in length (Supplementary Data 1). From this alignment, we identified eight conserved motifs based on >50% consensus across the MSA (Supplementary Data 1). Six of these motifs were able to be mapped to the structure of Red $\beta_{177}$  (Fig. 3c and Supplementary Fig. 4), the other two motifs are located in the C-terminal domain. The amino acid sequence conservation was also mapped onto the structure of Red $\beta_{177}$  (Fig. 3d). This showed that the most highly conserved residues, which are in motifs 2, 3, and 6, form the DNA-binding pocket. Two residues in particular, F66 and W143, are almost universally conserved across the sequences in the database.

### Oligomerization

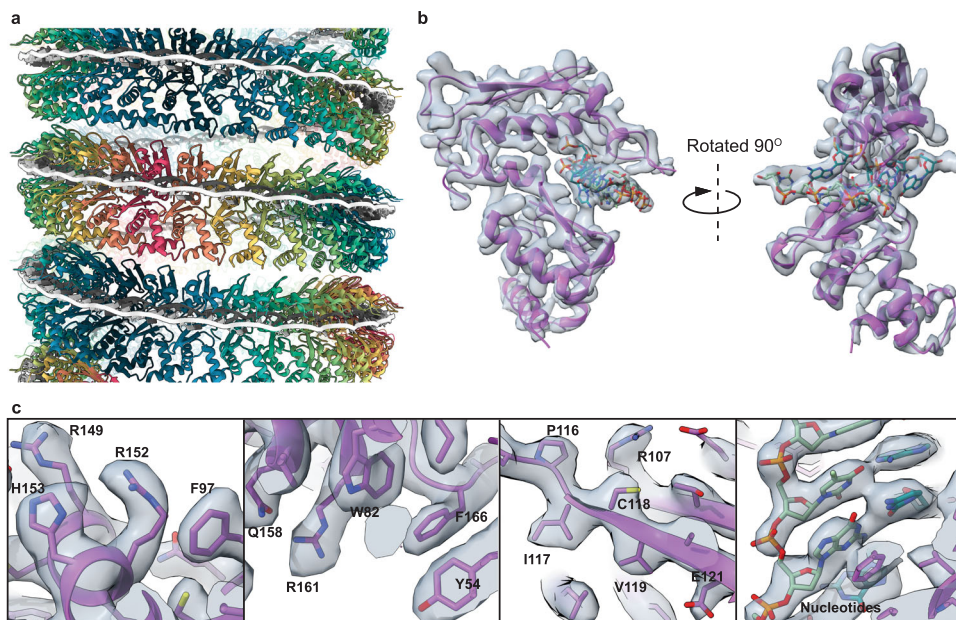
The model of Red $\beta_{177}$  was reconstructed as a compact, solenoid, helical filament with dsDNA wound around the outside. This superstructure is stabilized through substantial electrostatic contacts between adjacent monomers in the complex (Fig. 4a, b, Supplementary Fig. 5). Much of the large interaction surface area is comprised of patches of positive and negative charge, which are mirrored on the adjacent monomer face (Fig. 4b). In particular, a cluster of strong electrostatic interactions formed by E125, E121, K148', and R149'; R161 and D80'; D68, K61' and K36', may help to stabilize the area around the DNA-binding pocket (Fig. 4b). Some of these residues such as K36, K61, and R161 are also involved in DNA-binding. Therefore, oligomerization of adjacent units may also be indirectly facilitated by the binding of the DNA substrate, as described below.

The interacting bottom-to-top monomer surfaces contribute relatively little to the stabilization of the helix (Fig. 4c, d). There are only two points of contact, E102 and N101; and T107' and E14', between the top



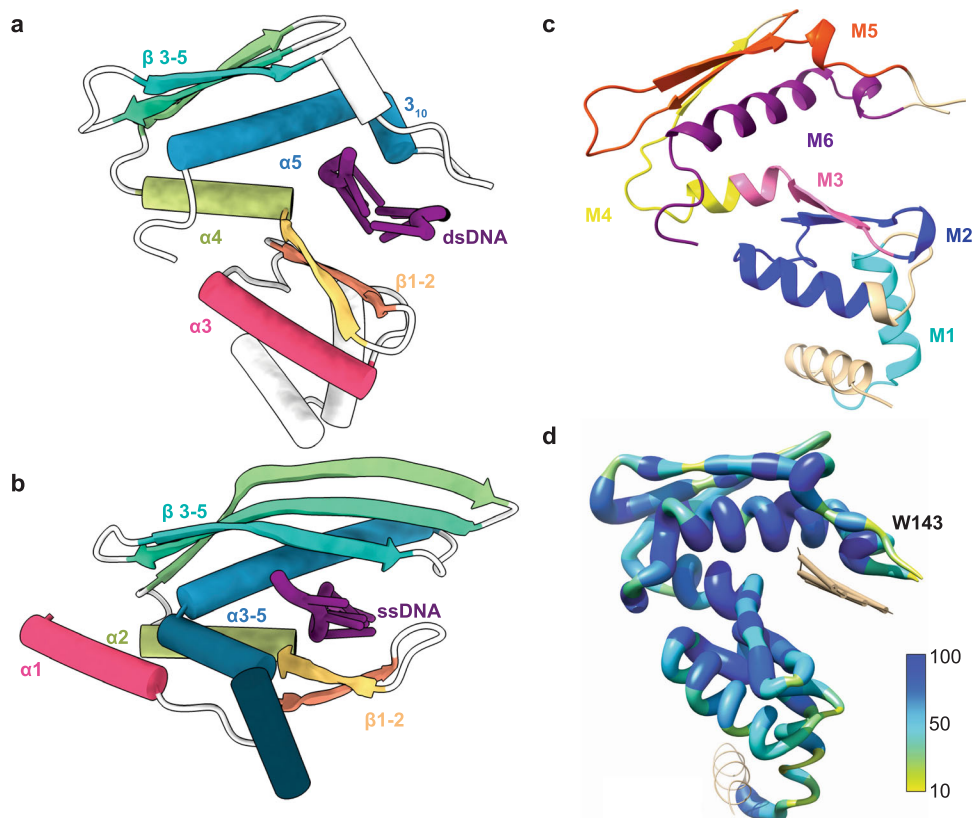
**Fig. 1 | Cryo-EM structure of Red $\beta_{177}$  1-start helical assembly, resolved to 3.3 Å.** **a** A representative electron micrograph displaying Red $\beta_{177}$  helices. Among the 4701 electron micrographs collected, all images identified as containing our protein sample displayed similar helical filaments. **b** A representative 2D class average of the helical filaments. **c** Fourier transform of a 2D class average. **d** Cryo-EM map of

the 1-start helical assembly resolved to 3.3 Å. **e** Local resolution of the structure visible from both a side and a top view. **f** Plot showing the Fourier shell correlation (FSC) versus spatial frequency, for both the masked (red) and unmasked (blue) maps. The resolution of the structure was assessed using the point where the FSC curve crosses a correlation value of 0.143.



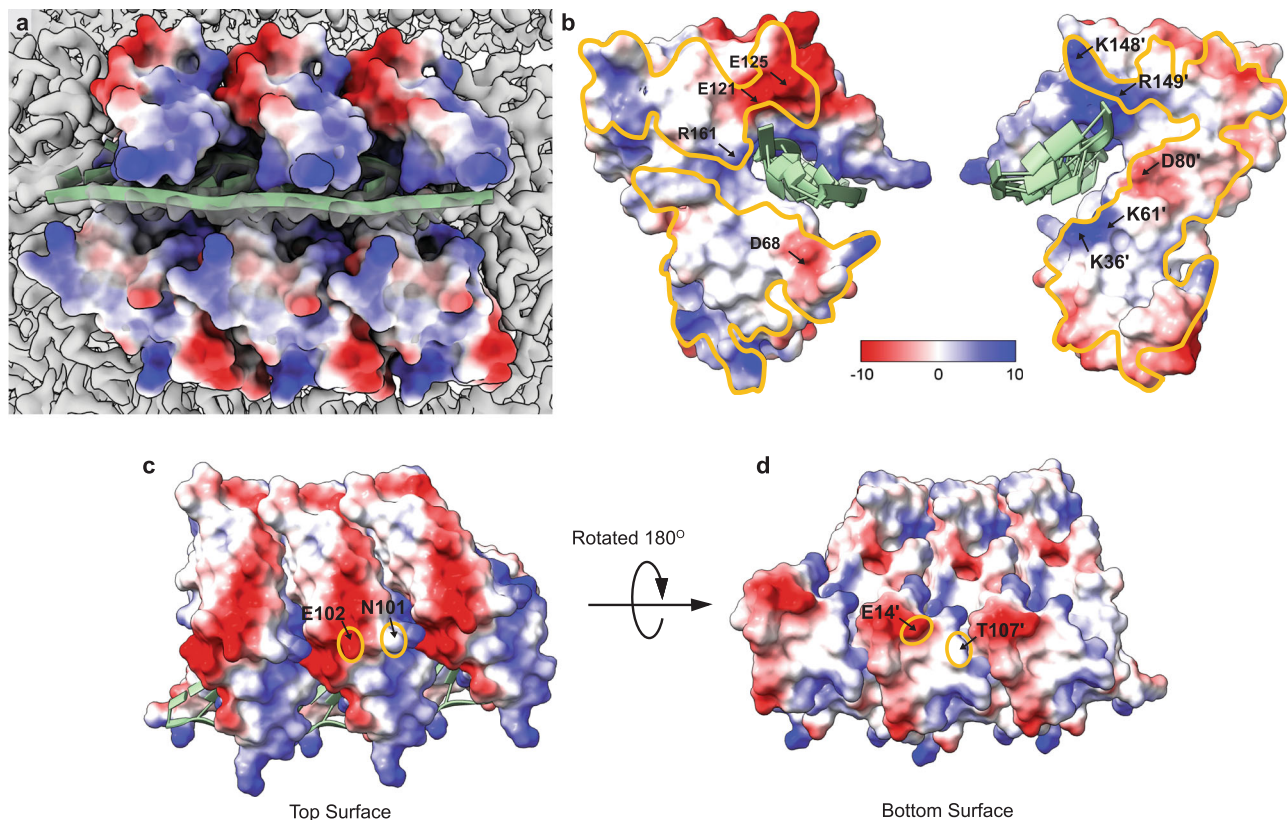
**Fig. 2 | Red $\beta_{177}$  and DNA structure.** **a** The ribbon structure of the Red $\beta_{177}$  helix showing the interaction between protein and DNA (black and white) in a planar conformation. **b** The cryo-EM map of a single Red $\beta_{177}$  monomer (gray) bound to

DNA, with the ribbon structure within it (purple). Close-up views of this monomer **c** demonstrates the fit of amino acids and nucleotides within the density map.



**Fig. 3 | Conserved annealase structural motifs.** Cartoon models comparing the secondary structures of Red $\beta_{177}$  (**a**) and Rad52 (5XRZ) (**b**). Color denotes conserved secondary structures between the proteins, while white represents regions that are not conserved between the structures and DNA is shown in purple. **c** The ribbon

structure of the Red $\beta_{177}$  monomer, with colors indicating conserved motifs 1–6 that were identified by the MSA (Supplementary Fig. 4). Beige indicates no conservation. **d** Amino acid sequence conservation of Red $\beta_{177}$  compared to other annealases. The extent of sequence conservation (%) is denoted by thickness and color.



**Fig. 4 | Electrostatic surface representation of Red $\beta_{177}$ .** The surface is colored by charge with positive (blue), negative (red) and electroneutral (white). **a** Electrostatic representation of three monomers Red $\beta_{177}$  bound to dsDNA (pale green). **b** The electrostatic interactions between two adjacent monomers within the

helix. The areas where the monomers interact are outlined in orange (**c**, **d**). Interactions between the top (**c**) and bottom (**d**) surfaces of the Red $\beta_{177}$  monomers within the helix. The orange ovals highlight the interacting residues.

and bottom surfaces, respectively. Therefore, the side-by-side contacts are the important determinants of helical characteristics such as the pitch and rise. Likewise, the abundance of electrostatic interactions between adjacent monomers drives elongation of the helical fragment.

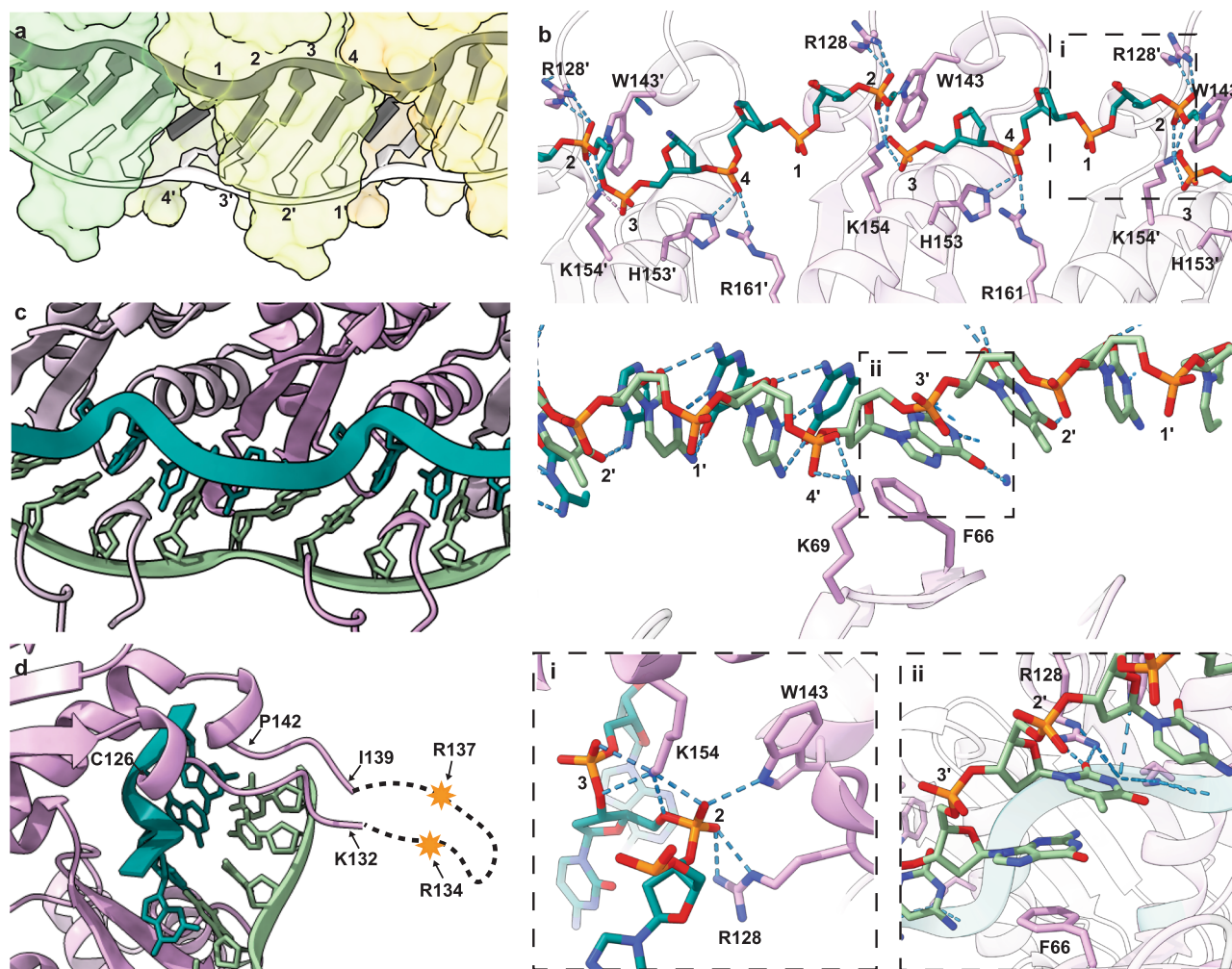
#### DNA-binding properties of Red $\beta_{177}$ and the mechanism for the catalysis of DNA annealing

In order to capture an annealing intermediate structure, which would shed light on the mechanism of SSA, our sample was prepared for cryo-EM by sequentially incubating Red $\beta_{177}$  with complementary ssDNA oligonucleotides. This approach has been previously used for Red $\beta$  in *in vitro* experiments to form annealing intermediates<sup>27,28,30</sup> and it resulted in clear density for a double-stranded DNA annealing intermediate in our structure. The outward-facing DNA-binding groove accommodates both strands of the annealed dsDNA product, with an observed stoichiometry of four base pairs per Red $\beta_{177}$  monomer (Fig. 5a). However, adjacent monomers bind the template strand in an overlapping manner, such that a single monomer makes contact with nucleotides involved in forming 6 base pairs. The DNA-binding groove is lined with positively charged residues (Fig. 4b), which accommodate the negatively charged phosphate backbone of the DNA, through a combination of electrostatic interactions and hydrogen bonds (Fig. 5b).

A total of nine hydrogen bonds between the phosphate backbone and the side chains of R128, W143, H153, K154 and R161 secure the inner, template ssDNA strand in the DNA-binding groove of Red $\beta_{177}$  (Fig. 5b, top panel). These bonds result in the nucleotides adopting a planar orientation and facilitate solvent exposure to provide a template for base pairing to incoming homologous

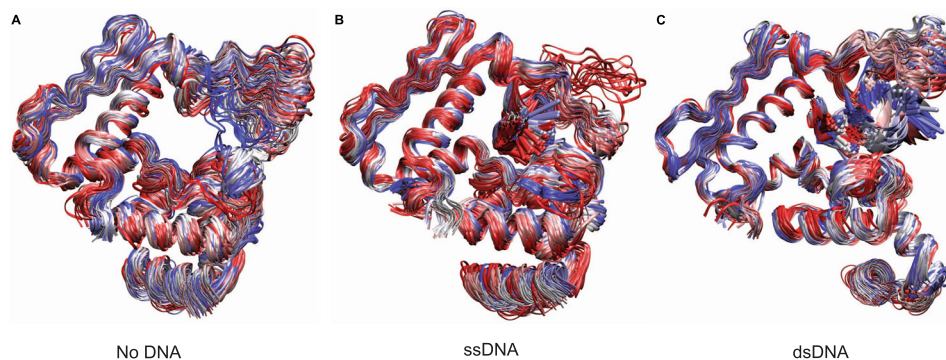
ssDNA. In contrast, the incoming DNA strand is only stabilized by two hydrogen bonds to the phosphate backbone through K69 (Fig. 5b bottom panel). In addition, the incoming strand is stabilized in a planar orientation through a  $\pi$ -stacking interaction between the phenyl ring of F66 and the purine or pyrimidine ring of the position 3' nucleotide and two hydrogen bonds between R128 and the adjacent nucleotide at position 2' (Fig. 5b, panel ii). These interactions may tether and orient the incoming ssDNA to the annealing complex, whilst allowing it to slide against the template strand until regions of base pairing stabilize the duplex. The annealed duplex DNA is also stabilized in an extended, planar conformation in the DNA-binding groove (Fig. 5), which is atypical of B-form helical dsDNA conformations often found in solution<sup>59</sup>. It is likely that upon annealing, the planar conformation is maintained through an induced distortion in the phosphate backbone of the inner (template) DNA strand, which is stabilized by a network of seven hydrogen bonds with R128, W143 and K154 (Fig. 5b, panel i). This kink (Fig. 5c) is necessitated due to the shorter helical radius, and thus shorter path of the inner DNA strand, relative to the outer strand.

Residues 133–138 could not be modeled in our structure. This region forms part of a larger, positively charged, loop (residues 126–142) that forms the roof of the DNA-binding groove (Fig. 5d). It is likely that residues 133–138 that form the tip of the loop are highly flexible, the motion of which (i.e., conformational variability) precluded the formation of structured density for this region during reconstruction. To see how this loop may be moving and functioning, we ran molecular dynamics (MD) simulations using a trimeric assembly of apo protein (Red $\beta_{177}$  only), Red $\beta_{177}$  bound either to template ssDNA or to both DNA strands. Molecular dynamics simulations were run



**Fig. 5 | DNA binding to Red $\beta_{177}$ .** dsDNA binds to Red $\beta_{177}$  with a site size of 4, as numbered in **a**, at the position of the phosphate of each nucleotide. **b** Hydrogen bonding between Red $\beta_{177}$ , the template (top panel) and incoming (bottom panel) DNA strands. Close-up views of the H-bonding network stabilizing the kink in the phosphate backbone (**i**) and the  $\pi$ -stacking interaction between F66 and a

nucleotide (**ii**) are shown. Ribbon diagrams of DNA (green) binding to three Red $\beta_{177}$  monomers (purple) clearly show the kink in the phosphate backbone of the template strand (**c**) as well as the finger loop (**d**), highlighting the residues which could not be modeled due to lack of density. A prime (') symbol denotes numbering in an adjacent molecule.



**Fig. 6 | Molecular dynamics simulations of Red $\beta_{177}$ .** The sampled ensemble of apo (**A**), ssDNA-bound (**B**) and dsDNA-bound (**C**) states used in the molecular dynamics simulations. Trimeric systems with weak harmonic positional restraints on the backbone atoms of two outside protein chains and DNA were simulated. One

hundred snapshots with a time interval of 10 ns from 1  $\mu$ s simulations are shown. The structures were fitted to the backbone atoms for the monomer in the middle of the trimeric system. Blue, silver, and red colors correspond to early, mid and late time intervals, respectively.

both with weak harmonic restraints applied on the two outside monomer and DNA backbone atoms (Fig. 6) and also without any restraints, as a control (Supplementary Fig. 6). Weak harmonic restraints help to maintain the oligomeric Red $\beta$  structure (modeled by

a trimeric structure) and the intermediate DNA structure. Each system was simulated for 1  $\mu$ s during which the RMSD of the protein backbone positions and RMSF of the C $\alpha$  atom of the central trimeric subunit were calculated (Supplementary Fig. 6d–g).

As we hypothesized, the MD results clearly demonstrate the flexibility of this loop in all simulations. However, the greatest range of movement, as indicated by the highest average RMSD, was seen in the absence of DNA (Fig. 6A and Supplementary Fig. 6a, d–g). The conformations of the loop varied from completely open, which would allow the diffusion of the incoming DNA strand into the binding site, to completely closed, which would keep the DNA in there. The range of motion decreases considerably in the presence of ssDNA and further still, in the presence of nascent dsDNA (Fig. 6B, C and Supplementary Fig. 6b–g). The decrease in motion is likely due to steric restrictions between the protein and the DNA. However, the flexible loop-tip contains two arginine residues, R134 and R137 (Fig. 5d), which may interact with the phosphate backbone of the incoming DNA strand to facilitate base pairing and thereafter helping to clamp the nascent dsDNA in place.

## Discussion

Despite its discovery in 1966<sup>60</sup>, the structure of Red $\beta$  annealase has remained unknown. It has proven to be extremely difficult to crystallize Red $\beta$ , despite trying a variety of truncations and added substrates. The crystals that could be grown did not diffract well enough to yield a structure (personal communication with Dr. Xinhua Ji at NCI/NIH, MD, USA and Dr. Charles Bell at OSU, OH, USA). Red $\beta_{177}$  used in this study was first described by the Bell group in 2006<sup>52</sup>. Our initial cryo-EM trials with apo protein yielded solenoid helices that were relaxed, and therefore, heterogeneous. Upon the addition of DNA and the optimization of buffer conditions, we have obtained compact solenoid helices of DNA-bound Red $\beta_{177}$  that are regular enough for helical image processing. The structure we have obtained unravels the molecular mechanisms of many observations and results published during the half a century since Red $\beta$  was discovered.

The first structural information of an annealase came from human Rad52<sup>48</sup>. Red $\beta$  and Rad52 share only ~17% amino acid sequence identity and previous bioinformatic studies placed the two proteins into superfamilies of their own<sup>7,47</sup>. However, similarity is evident at the level of the 3D structures (Fig. 3a, b) and the DALI server identified Rad52 as the most similar structure to Red $\beta_{177}$ . Structures of Rad52 bound to ssDNA have been reported in two different modes, described as ssDNA bound to either the inner or outer DNA-binding groove<sup>49</sup>. Comparing our structure of DNA-bound Red $\beta_{177}$  to the Rad52 structures, we see that the placement of the template ssDNA in Red $\beta_{177}$  is similar to the ssDNA found at the inner binding site of Rad52 (SXRZ). Both proteins have DNA-binding sites made up of the  $\beta$ - $\beta$ - $\alpha$ - $\beta$ - $\beta$ - $\alpha$  secondary structure motif that is predicted to be conserved among the annealases<sup>46</sup>. However, a major difference between the two proteins is the location of the flexible loop that is proposed to hold the DNA in place. In Red $\beta$ , the flexible loop forms the ‘top’ of the DNA-binding site, whilst in Rad52 a functionally similar loop is found at the ‘bottom’ of the DNA-binding site, located between the first two  $\beta$ -strands in the motif (Fig. 3a, b). The disparity between sequence and structure similarity between Red $\beta$  and Rad52 could be explained by two scenarios: (1) Both proteins are homologous and share a common ancestor but have diverged significantly over time. However, as structure is more conserved than sequence<sup>61</sup>, the proteins remained structurally similar despite divergent sequence evolution. (2) The proteins are not homologs, but their structure is biologically favorable for the role of an annealase. In this scenario, the similarities in protein structure would be the result of convergent evolution towards an enzymatic optimum.

DdrB from *Deinococcus radiodurans*, which was originally identified as belonging to the Rad52 superfamily of annealases<sup>7</sup>, has also been captured as a DNA annealing intermediate structure. Two pentameric DdrB rings assemble to form a DNA-binding channel bringing complementary ssDNA strands together facing each other in the middle of the double-ring complex<sup>62</sup>. Based on the DdrB structure,

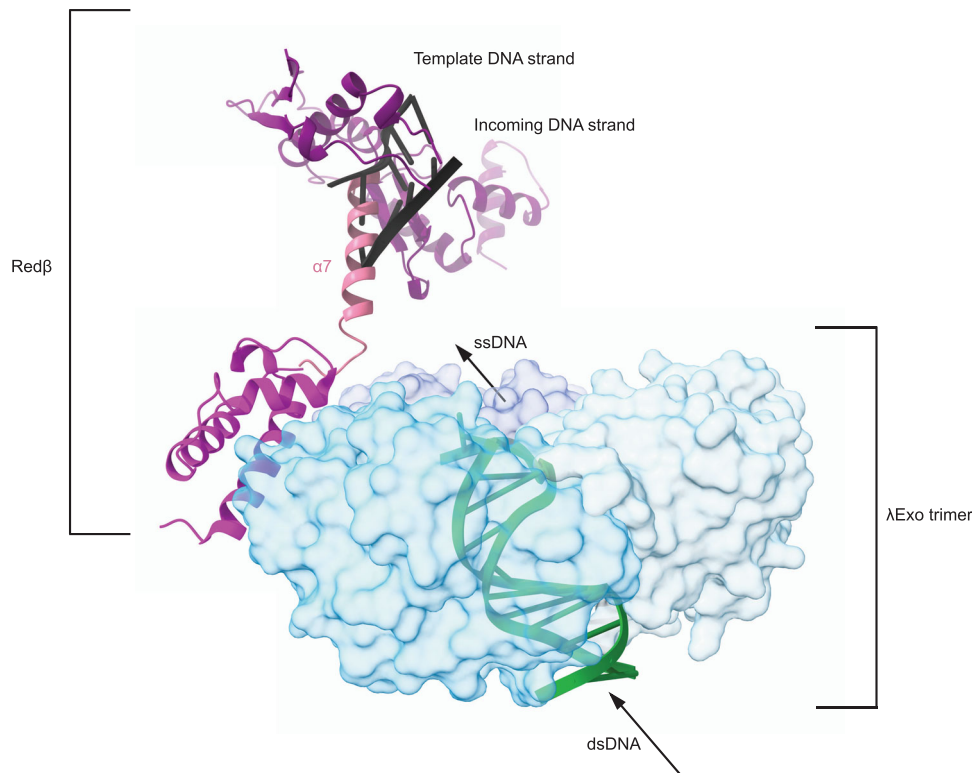
when we saw the conversion of the Red $\beta_{177}$  solenoid structure from relaxed to tight following the addition of complementary ssDNA strands, we predicted that ssDNAs would be bound in between the ‘rings’ of the helix (i.e., lock-washers forming each turn of the helical filament), with the DNA bases facing each other. We had attributed compaction of the solenoid to annealing between ssDNA strands, bringing the lock-washer rings together. Therefore, the location of the DNA in our reconstructed map, (wrapped around the outside of the helical Red $\beta_{177}$  filament) was an unexpected observation. A previous study proposed that ssDNA bound around the outside of Red $\beta$  rings but that annealed dsDNA was located on the inside of the helical filament<sup>27</sup>. We show that the predicted location of the ssDNA was correct, and that this DNA-binding groove is also the location of the annealed dsDNA. What is similar between the DdrB and Red $\beta$  is the conformation of the DNA in the annealing intermediate: the nascent dsDNA is found in a planar conformation, unlike the various helical conformations dsDNA forms in solution<sup>59</sup>. This also explains why Red $\beta$  does not bind to pre-annealed dsDNA<sup>26</sup>, as the double-helical structure known as the B-form DNA would not fit into the DNA-binding site of Red $\beta$ .

The homo-oligomerization of Red $\beta$  has been reported by several groups using various approaches and techniques, including native mass spectrometry<sup>30</sup>, electron microscopy<sup>27,50</sup> and atomic force microscopy<sup>28,29</sup>. These results are briefly summarized above and in more detail in our recent review<sup>8</sup>. Our reconstruction shows how Red $\beta_{177}$  oligomerizes with substantial contacts between adjacent subunits in the helix, but there are limited top-bottom contacts between corresponding monomers in the next turn of the helix (Fig. 4). This may explain the heterogeneity in the compactness of the solenoid-like helices of Red $\beta_{177}$  observed in the micrographs. Our structure also shows how DNA-binding reinforces interactions between adjacent monomers, favoring oligomerization (Fig. 5b). There are 4 bases per monomer in the DNA-binding groove, but each monomer makes contacts with the phosphates from 6 nucleotides of the template strand (Fig. 5a, b). This may explain the discrepancy in the literature regarding the site size of Red $\beta$ , which has been reported as 4–5 nts<sup>51,63</sup>.

Our structure has elucidated an extensive hydrogen bonding network involved in DNA binding, validating predictions made in previous studies. In particular, highly conserved R161 forms a hydrogen bond with position 4 of the template strand (Fig. 5b), supporting previous observations that mutation of this residue is correlated with a 1000-fold reduction in annealing efficiency *in vivo*<sup>46</sup>. Lysines at positions 36, 61, 111, 132, 148, 154, and 172 were reported to be critical for DNA binding *in vitro* and recombination *in vivo*. K61 was specifically reported to be critical for DNA annealing but not for initial ssDNA binding, and a role in binding to the incoming strand of DNA was suggested<sup>64</sup>. Our structure shows that lysines 36, 61, 132, and 154 are located close to the DNA-binding site, but 111 is not (K172 is unresolved in our structure). In addition, K69, which was not indicated to be involved in DNA-binding previously, is also shown to interact with the DNA (Fig. 5b). Interestingly, K61 is closer to the template DNA than to the incoming DNA strand, while lysines 36, 69, 132 are closer to the incoming strand. Although K148 is within close proximity to the DNA-binding site, its side chain points away from the DNA. Moreover, other positively charged residues H146, H153, R128 and R149 are also near the DNA, with R128 located within hydrogen bonding distance of both DNA strands, while W143 can make a hydrogen bond with the template strand.

Red $\beta$  catalyzed annealing of homologous ssDNA was reported to be an apparent first-order reaction (the half-time of the reaction was independent of the DNA concentration) and a synaptic mechanism was suggested in which the DNA molecules are first brought together irrespective of homology, which is then followed by homology searching and base pairing<sup>63</sup>. The annealing intermediate structure of





**Fig. 7 | Superimposed molecular structures of Red $\beta$ ,  $\lambda$ Exo and DNA.** This molecular scene was created by piecing four structures together. These include the structure of  $\lambda$ Exo bound to DNA<sup>66</sup> (4WUZ, three  $\lambda$ Exo monomers represented as surfaces in different shades of blue and DNA in green); the structure of  $\lambda$ Exo bound to the C-terminal domain of Red $\beta$ <sup>53</sup> (6M9K, C-terminal domain of Red $\beta$  shown in

light purple and  $\lambda$ Exo superimposing with the one in 4WUZ is hidden); the Red $\beta$ <sub>177</sub> model (DNA in black and protein in dark purple); and the AlphaFold prediction for the full-length Red $\beta$ <sup>54</sup> (the  $\alpha$ 7 helix shown in pink and the N- and C-terminal domains superimposing with the corresponding domains in our Red $\beta$ <sub>177</sub> structure and 6M9K, respectively, are hidden).

Red $\beta$ <sub>177</sub> that we determined reveals the molecular mechanism for a first-order annealing reaction. Our sample was generated by the sequential addition of the two complementary 27 base pair oligonucleotides to Red $\beta$ <sub>177</sub> and the structure presented suggests that the binding of the template strand to Red $\beta$  is required as the initial step. After this, we hypothesized that the following events may be taking place: the incoming complementary ssDNA can bind transiently while base pairing (and therefore, homology) between the two DNA strands is being evaluated, possibly held in place by the flexible loop (composed of residues 130–139 most of which are unresolved in our structure) to which we refer to as the ‘finger loop’. Once the finger loop is reopened, the incoming ssDNA would stay in place if a strong enough base pairing is achieved; otherwise, it would diffuse out or slide to allow a new base pair match to be tested. This opening-closing and 2D and/or 3D diffusion of the incoming ssDNA would be iterated until homologous pairing is established between the two ssDNA strands bound to Red $\beta$ .

To test our hypothesis for the action of the finger loop, we ran molecular dynamics simulations. Our results showed that the mechanism we proposed is supported by the finger loop moving between open and closed conformations. The motion of the finger loop is likely a dynamic process that can be described as tapping behind the incoming ssDNA, keeping it in place to allow more time for the homology search to take place, but loose enough to allow sliding (2D diffusion) and thereby the detection of homology in situations where the nucleotide sequence may be shifted. The simulations also showed that Red $\beta$  is most flexible when it is not bound to any DNA and then becomes more rigid when bound to the template ssDNA. Red $\beta$  is then further stabilized upon forming an annealing complex with two DNA strands. This is in agreement with the reports of Red $\beta$  forming stable DNA annealing complexes<sup>26</sup>.

Red $\beta$ <sub>177</sub> contains a  $3_{10}$  helix between the finger loop and  $\alpha$ 5;  $3_{10}$  helices are often located between an alpha helix and a flexible region, where the weaker  $i + 3$  H-bonding network is proposed to allow transient unraveling<sup>65</sup>. Along with  $\alpha$ 5 this  $3_{10}$  helix forms part of motif 6 (Fig. 3c and Supplementary Fig. 4) which is highly conserved across annealases (Supplementary Data 1). In addition, W143 of the  $3_{10}$  helix is almost universally conserved across all species. Our structure identified that W143 forms part of the H-bonding network that stabilizes the induced kink in the phosphate backbone of the template DNA strand. Therefore, the position of W143 in this  $3_{10}$  helix is likely to be of functional importance. The MD simulations indicated that the  $3_{10}$  helix was unstable, oscillating through distorted conformations in concert with the movement of the finger loop. Therefore, we propose a mechanism where transient relaxation of the  $3_{10}$  helix as the finger loop opens allows H-bonding of W143 to the phosphate backbone of the template strand. Then, as the loop closes and the  $3_{10}$  helix reforms, the phosphate backbone may be pulled into the distorted conformation seen in our structure. The distortion in the phosphate backbone allows the template DNA strand to remain planar. This facilitates base pairing between the nucleotides of the template and incoming strands, both of which are held in a planar orientation through F66 ( $\pi$ -stacking) and R128 (H-bond), enhancing the formation of a stable annealing intermediate.

Finally, we have created a molecular scene by piecing four structures together – 4WUZ ( $\lambda$ Exo with DNA<sup>66</sup>), 6M9K ( $\lambda$ Exo with the C-terminal domain of Red $\beta$ <sup>53</sup>), our Red $\beta$ <sub>177</sub> model, and the AlphaFold prediction (Red $\beta$ <sub>AF</sub><sup>54</sup>) of full-length Red $\beta$  (Supplementary Fig. 7). This scene (Fig. 7) shows dsDNA bound to  $\lambda$ Exo, from which the digested nascent ssDNA would be extruded from the center of the  $\lambda$ Exo toroid towards the Red $\beta$  molecules bound to the ‘back’ side of  $\lambda$ Exo (only one Red $\beta$  is shown, but there would be 3, one for each  $\lambda$ Exo monomer). The

Red $\beta_{AF}$  prediction shows helix  $\alpha 7$  placed near the DNA-binding pocket of Red $\beta$ . Therefore, the superposition of the Red $\beta_{177}$  and Red $\beta_{AF}$  places  $\alpha 7$  in a position that is partially clashing with the bound DNA. If AlphaFold is correctly predicting the position of this helix in the apo state, our structure of Red $\beta$  shows that helix  $\alpha 7$  must be displaced via a conformational change to allow DNA binding. This may function as an auto-inhibition mechanism, to ensure that Red $\beta$  would not bind to ssDNAs generated in the cell during other processes, such as DNA replication, and that it is only able to bind to nascent ssDNA generated via digestion of dsDNA by  $\lambda$ Exo. The nascent ssDNA would be effectively “loaded” onto Red $\beta$  by this coupling mechanism, coordinating the two reactions catalyzed by the individual components of this two-component recombinase: DNA digestion by  $\lambda$ Exo and DNA binding and annealing by Red $\beta$ . Our molecular scene provides a glimpse of what this coupling may look like, furthering our understanding of the molecular mechanisms of SSA.

There are many questions that remain about the molecular mechanisms of SSA: how can a trimeric  $\lambda$ Exo (3 subunits) and a dodecameric Red $\beta$  (12 subunits) form a complex with 1:1 monomer:monomer stoichiometry? Or, how does Red $\beta$  form rings and helices on ssDNA in concert with dsDNA digestion by  $\lambda$ Exo? Moreover, the structure we determined poses questions, such as what is the function and significance of the finger loop in Red $\beta$  and Rad52 and how is the nascent dsDNA released from the stable Red $\beta$  helices? As efforts to answer those questions are ongoing, our structure of Red $\beta_{177}$  provides insights into the mechanism of how this protein functions.

## Methods

### Cloning of Red $\beta_{177}$

A pET-28a(+) plasmid containing the wild-type *bet* gene was kindly provided by the laboratory of Dr. Donald Court (National Cancer Institute, NIH, Frederick, MD, USA). Red $\beta_{177}$  was assembled into a pET-24a(+) plasmid for expression with a C-terminal hexa-histidine (His $_6$ ) tag. The DNA sequence corresponding to Red $\beta_{177}$  was PCR amplified using primers 5'-GTTTAACCTTAAGAAGGAGATATACCATGAGTACTGCACTCGCAAC-3' and 5'-GATCTCAGTGGTGGTGGTGGTGGTGAGAA-GAGCGCTCGGCTTCATCCTTG-3', for the simultaneous addition of pET-24a(+) overlap sequences. The pET-24a(+) plasmid was PCR amplified using primers 5'-GGTATATCTCCTTCTAAAGTTAAACAAA TTATTTCTAG-3' and 5'-TCTTCTACCACCACCACCACCAGTATC-3'. Following PCR, both DNA fragments were purified by agarose gel electrophoresis and subsequent extraction. The recombinant plasmid was generated by a NEBuilder<sup>®</sup> HiFi DNA Assembly (New England Biolabs) reaction, under standard conditions and using a 2:1 molar ratio of insert:plasmid DNA. The resulting construct, pET-24a-bet $_{177}$ -His $_6$ , places the expression of Red $\beta_{177}$ -His $_6$  (Red $\beta_{177}$ ) under the control of the *E. coli lac* expression system. The correct sequence of the plasmid was confirmed by Sanger sequencing (Garvan Molecular Genetics, Garvan Institute of Medical Research, Sydney, Australia). *E. coli* strain BL21(DE3) was transformed by the recombinant plasmid and used for protein expression.

### Expression and purification of Red $\beta_{177}$ protein

*E. coli* BL21(DE3) cells were cultured in LB media supplemented with 50  $\mu$ g/ml kanamycin, at 37 °C and with shaking at 200 rpm. When the culture reached an OD $_{600\text{ nm}}$  of 0.6, protein expression was induced by the addition of 1 mM isopropyl- $\beta$ -D-1-thiogalactopyranoside (IPTG, Astral Scientific). At 3 h post induction, the cells were harvested by centrifugation at 6100  $\times g$  for 7 min at 6 °C. The cell pellet was resuspended in 50 ml of Lysis buffer (50 mM Tris pH 7.5, 300 mM NaCl, 10% *v/v* glycerol, 10% *w/v* sucrose) and cells were lysed in an ice bath by sonic pulses produced by an Ultrasonics digital sonifier (Branson). The lysate was clarified by centrifugation at 38,000  $\times g$  for 30 min at 4 °C. Following centrifugation, the supernatant was passed through a 0.45  $\mu$ m syringe-driven filter (Sarstedt).

At a flow rate of 1 ml/min, controlled by an ÄKTA Pure FPLC equipped with a sample pump (Cytiva), the filtrate was pumped through a 1 ml HisTrap HP column (Cytiva), pre-equilibrated with 50 mM Tris pH 7.5, 300 mM NaCl, 10% *v/v* glycerol and 10 mM imidazole. The flow-through from the column, containing unbound proteins, was collected and retained. The column was washed at 1 ml/min with 10 column volumes of wash buffer (50 mM Tris pH 7.5, 300 mM NaCl, 10% *v/v* glycerol and 20 mM imidazole), before bound proteins were eluted with a 0–100% gradient of elution buffer (50 mM Tris pH 7.5, 300 mM NaCl, 10% *v/v* glycerol and 500 mM imidazole), over 30 column volumes at a flow rate of 1 ml/min. The column was then washed with 10 column volumes of elution buffer, before being re-equilibrated.

The flow-through from the first separation was pumped back onto the column for further extraction of Red $\beta_{177}$ . Elution fractions from both chromatography separations were analyzed by SDS-PAGE. Fractions that were highly enriched for Red $\beta_{177}$  were pooled and concentrated in a 10 kDa MWCO centrifugal concentrator (Merck Millipore) to a total volume of <2 ml. The concentrated protein was centrifuged at 21,000  $\times g$ , for 30 min at 4 °C, following which there was no visible pellet. The supernatant was injected onto a HiLoad 16/600 Superdex pg gel filtration column (Cytiva), equilibrated with 50 mM Tris pH 7.5, 50 mM NaCl, 10% *v/v* glycerol. Red $\beta_{177}$  eluted as a single, symmetric peak at 1 ml/min flow. The central peak fractions were pooled and estimated to be >95% pure by SDS-PAGE. Protein concentration was assessed by measuring absorbance at 280 nm using a NanoDrop 2000c spectrophotometer (ThermoFisher Scientific) and the extinction coefficient of  $\epsilon = 34,950 \text{ L mol}^{-1} \text{ cm}^{-1}$ . Protein was stored stable for up to 3 weeks at 4 °C, or flash frozen in liquid nitrogen and then stored at -80 °C.

### Helical complex assembly and preparation of cryo-EM grids

To form the helical assemblies, purified Red $\beta_{177}$  was incubated with complementary 27mer oligonucleotides. Red $\beta_{177}$  was incubated with Oligo 1 (5'-TGCAGCAGCTTTACCATCTGCCGCTGG-3') in binding buffer (20 mM KH $_2$ PO $_4$  pH 6.0, 5 mM MgCl $_2$ ) for 30 min at room temperature. Oligo 2 (5'-CCAGCGGCAGATGGTAAAGCTGCTGCA-3') was then added, and the incubation was continued for a further 30 minutes. A molar ratio of 7.7:1:1 of Red $\beta_{177}$ :Oligo 1:Oligo 2 was used. To remove any remaining traces of glycerol from the storage buffer, the complex was dialyzed against the binding buffer overnight at 4 °C, using a 10 kDa MWCO Slide-A-Lyzer MINI Dialysis cassette (ThermoFisher Scientific). A 2.5  $\mu$ l aliquot of the dialyzed complex was deposited onto gold UltrAuFoil RL2/L3 Au 300 mesh cryo-electron microscopy grids (Quantifoil) and plunge frozen into liquid-nitrogen cooled liquid ethane using a Mark IV Vitrobot (ThermoFisher Scientific).

### Cryo-electron microscopy and helical reconstruction of Red $\beta_{177}$ -6xHis helical assemblies

The data were collected using a Titan Krios electron microscope (ThermoFisher Scientific) which operated at 300 kV. The Titan Krios was equipped with a Gatan K2 Summit detector and a Gatan Bio-Quantum LS 967 energy filter, operated in unfiltered mode. Initial microscope alignment was performed using TEM User Interface (ThermoFisher Scientific) and DigitalMicrograph (Gatan). Data were collected in electron counting mode, using a pixel size of 0.84 Å/pixel as well as a calibrated sample-to-pixel magnification of  $\times 59524$ .

Movies were collected as a series of 50 frames with a total accumulated dose of 50 e $^{-}/\text{Å}^2$ . Data were collected using automated data collection in EPU, with a defocus range of -0.6 to -2.2  $\mu$ m. All processing was performed in cryoSPARC<sup>55</sup> unless otherwise noted. Movies were gain corrected and aligned using patch-based motion correction and patch-based CTF correction was performed. Helical filament tracing and extraction was performed with a box size of 576 pixels and a particle separation distance of 55 pixels. Extracted particles were

subjected to 2D classification and subsequent 2D class averages were used as templates for further rounds of template-based helical filament tracing. This process was repeated twice until picking had reached sufficient coverage of the helical filaments. Particles were then re-extracted with a box size of 448 pixels and rescaled to a box size of 224 pixels to assist with the speed of reconstruction. Extracted particles were subjected to several rounds of 2D classification to identify suitable subsets for reconstruction. Fast Fourier transforms of 2D classes were generated in Fiji<sup>67</sup>.

Suitable classes identified via 2D classification were initially subjected to 3D helical reconstruction with no helical symmetry applied. Once helical reconstruction had converged, particles were re-extracted to a box size of 448 pixels with no rescaling. Helical reconstruction was then repeated, and the helical symmetry parameters were estimated using real-space helical symmetry estimation. Helical reconstruction was then performed with helical symmetry and subsequent non-uniform refinement. CTF refinement and local motion correction were then performed until convergence. The final resolution of the reconstructed density was estimated using the gold-standard Fourier shell correlation (GSFSC) criteria.

### Structure building and refinement

De novo atomic modeling in COOT<sup>68</sup> was used to generate the initial structure for the Red $\beta$ <sub>177</sub> monomer and the bound nucleic acid. The model generated in COOT was then refined using ISOLDE<sup>69</sup>. As the density corresponding to the bases of the bound oligonucleotides within the asymmetric unit represents a helically symmetrized average density of the base pairs within the oligonucleotide substrate sequence, insufficient detail was available to assign a specific base identity to each of the positions in the asymmetric unit or to reliably model the associated hydrogen bonding between strands. To model the nucleic acids in the asymmetric unit, an arbitrary representative sequence from the substrate oligonucleotides was selected and standard Watson-Crick base pairing applied. The atypical confirmation of the template and non-template strands was modeled with the assistance of adaptive distance restraints in ISOLDE. Model statistics were assessed with MolProbity<sup>70</sup> as implemented in the Phenix software suite<sup>71</sup>.

### Bioinformatics

In order to thoroughly search the available sequence data for homologous sequences, a full-length Red $\beta$  amino acid sequence (UniProtKB: P03698) was BLAST searched ( $E$ -value cut-off = 1, substitution matrix = BLOSUM-64) against the UniProt Reference Cluster (UniRef) 90% database<sup>72</sup>. From this search, the top 1000 hits were extracted, and amino acid sequences aligned with the command line version of MAFFT (Multiple Alignment using Fast Fourier Transform) v.7.480<sup>73</sup> using the L-INSI algorithm. All MAFFT parameters were left at default, except for the 'maxiterate' parameter that was set to 1000.

The resultant multiple sequence alignment (MSA) was then visualized within Jalview v.2.11.1.4<sup>74</sup> and trimmed to the region of Red $\beta$  containing the residues 19–205. Redundant sequences and sequences less than 100 residues were then removed. Outlier sequences were also identified using the principal component and neighbour-joining tree analysis tools within Jalview (both using the default settings) and removed. Following refinement, the MSA was then again aligned with MAFFT using the above parameters.

Conserved motifs were then identified based on 50% consensus across the MSA as calculated in Geneious Prime v.2021.1.1 (available from: <https://www.geneious.com/>). Conserved motifs were then manually refined by identifying, and subsequently annotating, positions at which an amino acid property, rather than an individual residue, was conserved across the MSA.

DALI protein structure comparison server<sup>58</sup> was used with the default parameters for identifying the proteins that are structurally similar to Red $\beta$ <sub>177</sub>.

### Molecular dynamics simulations and structure prediction

Molecular dynamics simulations were carried out using NAMD (Nanoscale Molecular Dynamics, version 3 alpha)<sup>75</sup>. In total, three systems have been simulated each for 1  $\mu$ s, including Red $\beta$  alone (apo), Red $\beta$  bound to ssDNA (ss), and Red $\beta$  bound to the dsDNA annealing intermediate (ds). For each system, two different setups have been simulated: one set of the simulations were free simulations, while a positional harmonic restraint on the backbone atoms of DNA and two outside proteins was applied for the other set of simulations. The latter was used to probe the effects of using a trimeric structure to study the oligomeric assembly. Initial preparation of systems was done with CHARMM-GUI<sup>76,77</sup>. The unresolved loops (residue 131–136) were built with Modeller (version 10.1)<sup>78</sup>. The AMBER protein parm14SB force field<sup>79</sup> and nucleic acid BSC1 force field<sup>79</sup> were applied for the protein and DNA respectively. The TIP3P model was used for water<sup>80</sup>.

Molecular dynamics simulations were performed after solvating the system in an  $-100 \times 100 \times 100$  Å cubic box that extended at least 10 Å from the solute surface. Na<sup>+</sup> and Cl<sup>-</sup> counter ions were added to neutralize the system and achieve a salt concentration of 0.15 M. pKa calculations were performed using PROPKA to assign protonation states of ionizable residues. Simulations were performed using periodic boundary conditions (PBC) at constant temperature (303.15 K) with the Langevin algorithm (a damping coefficient of 1/ps)<sup>81</sup> and at a pressure of 1.0 bar using the Nose-Hoover Langevin Piston method<sup>82</sup>. The time step was set to 2.0 fs with all covalent bonds involving hydrogens kept rigid with the RATTLE algorithm<sup>83</sup>. Short-range electrostatics were calculated together with long-range electrostatics particle mesh Ewald (PME)<sup>84</sup> with a cut-off of 9.0 Å and a PME grid size of 1.0 Å. For all systems, energy minimization (10,000 steps) and 125 ps equilibration were performed first with positional restraints placed on all the protein and DNA heavy atoms (with a force constant of 1.0 kcal/mol/Å<sup>2</sup> on the backbone atoms and 0.5 kcal/mol/Å<sup>2</sup> on the side chain atoms). This was followed by 1  $\mu$ s production runs. In the simulation set with a positional harmonic restraint, this restraint was maintained on the two outside monomers and DNA backbones. Snapshots were saved every 100 ps. VMD (Visual Molecular Dynamics)<sup>85</sup>, LOOS (Lightweight Object-Oriented Structure Analysis)<sup>86</sup> and MDAnalysis<sup>87</sup> were used to analyze the trajectories.

Structural prediction was performed with the ColabFold<sup>88</sup> implementation of AlphaFold2. The source code was retrieved from the GitHub repository (<https://github.com/YoshitakaMo/localcolabfold>). For each prediction, five models were generated and optimized with the AMBER force field.

### Reporting summary

Further information on research design is available in the Nature Research Reporting Summary linked to this article.

### Data availability

The data that support this study are available from the corresponding author upon request. The coordinates for the atomic model have been deposited into the Protein Data Bank with the accession code 7UJL [<https://doi.org/10.2210/pdb7UJL/pdb>]. The cryo-EM map has been deposited into the Electron Microscopy Data Bank with the accession code EMD-26566. The sequence data used in the analysis described in this manuscript are publicly available in the UniProt Reference Cluster 90% database [<https://www.uniprot.org/uniref?facets=identity%3A0.9&query=%2A>].

### References

1. Murphy, K. C. Phage recombinases and their applications. *Adv. Virus Res.* **83**, 367–414 (2012).
2. Campbell, A. Comparative molecular biology of lambdoid phages. *Annu. Rev. Microbiol.* **48**, 193–222 (1994).

3. Piazza, A. & Heyer, W. D. Homologous recombination and the formation of complex genomic rearrangements. *Trends Cell Biol.* **29**, 135–149 (2019).
4. Weller, S. K. & Sawitzke, J. A. Recombination promoted by DNA viruses: phage lambda to herpes simplex virus. *Annu. Rev. Microbiol.* **68**, 237–258 (2014).
5. Bhargava, R., Onyango, D. O. & Stark, J. M. Regulation of single-strand annealing and its role in genome maintenance. *Trends Genet.* **32**, 566–575 (2016).
6. Enquist, L. W. & Skalka, A. Replication of bacteriophage lambda DNA dependent on the function of host and viral genes. I. Interaction of red, gam and rec. *J. Mol. Biol.* **75**, 185–212 (1973).
7. Iyer, L. M., Koonin, E. V. & Aravind, L. Classification and evolutionary history of the single-strand annealing proteins, RecT, Redbeta, ERF and RAD52. *BMC Genomics* **3**, 8 (2002).
8. Brewster, J. L. & Tolun, G. Half a century of bacteriophage lambda recombinase: In vitro studies of lambda exonuclease and Red-beta annealase. *IUBMB Life* **72**, 1622–1633 (2020).
9. Stahl, M. M. et al. Annealing vs. invasion in phage lambda recombination. *Genetics* **147**, 961–977 (1997).
10. Ellis, H. M., Yu, D., DiTizio, T. & Court, D. L. High efficiency mutagenesis, repair, and engineering of chromosomal DNA using single-stranded oligonucleotides. *Proc. Natl Acad. Sci. USA* **98**, 6742–6746 (2001).
11. Zhang, J., McCabe, K. A. & Bell, C. E. Crystal structures of lambda exonuclease in complex with DNA suggest an electrostatic ratchet mechanism for processivity. *Proc. Natl Acad. Sci. USA* **108**, 11872–11877 (2011).
12. Copeland, N. G., Jenkins, N. A. & Court, D. L. Recombineering: a powerful new tool for mouse functional genomics. *Nat. Rev. Genet.* **2**, 769–779 (2001).
13. Muylers, J. P., Zhang, Y. & Stewart, A. F. Techniques: recombinogenic engineering—new options for cloning and manipulating DNA. *Trends Biochem. Sci.* **26**, 325–331 (2001).
14. Kmiec, E. & Holloman, W. K. Beta protein of bacteriophage lambda promotes renaturation of DNA. *J. Biol. Chem.* **256**, 12636–12639 (1981).
15. Caldwell, B. J. & Bell, C. E. Structure and mechanism of the Red recombination system of bacteriophage lambda. *Prog. Biophys. Mol. Biol.* **147**, 33–46 (2019).
16. Marsic, N., Roje, S., Stojiljkovic, I., Salaj-Smic, E. & Trgovcevic, Z. In vivo studies on the interaction of RecBCD enzyme and lambda Gam protein. *J. Bacteriol.* **175**, 4738–4743 (1993).
17. Kulkarni, S. K. & Stahl, F. W. Interaction between the Sbc gene of *Escherichia coli* and the gam gene of phage-lambda. *Genetics* **123**, 249–253 (1989).
18. Lin, F. L., Sperle, K. & Sternberg, N. Model for homologous recombination during transfer of DNA into mouse L cells: role for DNA ends in the recombination process. *Mol. Cell Biol.* **4**, 1020–1034 (1984).
19. Poteete, A. R. Involvement of DNA replication in phage lambda Red-mediated homologous recombination. *Mol. Microbiol.* **68**, 66–74 (2008).
20. Maresca, M. et al. Single-stranded heteroduplex intermediates in lambda Red homologous recombination. *BMC Mol. Biol.* **11**, 54 (2010).
21. Mosberg, J. A., Lajoie, M. J. & Church, G. M. Lambda red recombineering in *Escherichia coli* occurs through a fully single-stranded intermediate. *Genetics* **186**, 791–799 (2010).
22. Wang, H. H. et al. Programming cells by multiplex genome engineering and accelerated evolution. *Nature* **460**, 894–898 (2009).
23. McEntee, K., Weinstock, G. M. & Lehman, I. R. recA protein-catalyzed strand assimilation: stimulation by *Escherichia coli* single-stranded DNA-binding protein. *Proc. Natl Acad. Sci. USA* **77**, 857–861 (1980).
24. Shinohara, A., Shinohara, M., Ohta, T., Matsuda, S. & Ogawa, T. Rad52 forms ring structures and co-operates with RPA in single-strand DNA annealing. *Genes Cells* **3**, 145–156 (1998).
25. Antony, E. & Lohman, T. M. Dynamics of *E. coli* single stranded DNA binding (SSB) protein-DNA complexes. *Semin. Cell Dev. Biol.* **86**, 102–111 (2019).
26. Karakousis, G. et al. The beta protein of phage lambda binds preferentially to an intermediate in DNA renaturation. *J. Mol. Biol.* **276**, 721–731 (1998).
27. Passy, S. I., Yu, X., Li, Z., Radding, C. M. & Egelman, E. H. Rings and filaments of beta protein from bacteriophage lambda suggest a superfamily of recombination proteins. *Proc. Natl Acad. Sci. USA* **96**, 4279–4284 (1999).
28. Erler, A. et al. Conformational adaptability of red beta during DNA annealing and implications for its structural relationship with Rad52. *J. Mol. Biol.* **391**, 586–598 (2009).
29. Subramaniam, S. et al. DNA annealing by Redbeta is insufficient for homologous recombination and the additional requirements involve intra- and inter-molecular interactions. *Sci. Rep.* **6**, 34525 (2016).
30. Caldwell, B. J. et al. Oligomeric complexes formed by Redbeta single strand annealing protein in its different DNA bound states. *Nucleic Acids Res.* **49**, 3441–3460 (2021).
31. Makhov, A. M. & Griffith, J. D. Visualization of the annealing of complementary single-stranded DNA catalyzed by the herpes simplex virus type 1 ICP8 SSB/recombinase. *J. Mol. Biol.* **355**, 911–922 (2006).
32. Tolun, G., Makhov, A. M., Ludtke, S. J. & Griffith, J. D. Details of ssDNA annealing revealed by an HSV-1 ICP8-ssDNA binary complex. *Nucleic Acids Res.* **41**, 5927–5937 (2013).
33. Ozgur, S., Damania, B. & Griffith, J. The Kaposi's sarcoma-associated herpesvirus ORF6 DNA binding protein forms long DNA-free helical protein filaments. *J. Struct. Biol.* **174**, 37–43 (2011).
34. Van Dyck, E., Hajjibagheri, N. M. A., Stasiak, A. & West, S. C. Visualization of human Rad52 protein and its complexes with hRad51 and DNA. *J. Mol. Biol.* **284**, 1027–1038 (1998).
35. Singleton, M. R., Wentzell, L. M., Liu, Y., West, S. C. & Wigley, D. B. Structure of the single-strand annealing domain of human RAD52 protein. *Proc. Natl Acad. Sci. USA* **99**, 13492–13497 (2002).
36. Ranatunga, W. et al. Human RAD52 exhibits two modes of self-association. *J. Biol. Chem.* **276**, 15876–15880 (2001).
37. Stasiak, A. Z. et al. The human Rad52 protein exists as a heptameric ring. *Curr. Biol.* **10**, 337–340 (2000).
38. Ruyechan, W. T. & Weir, A. C. Interaction with nucleic acids and stimulation of the viral DNA polymerase by the herpes simplex virus type 1 major DNA-binding protein. *J. Virol.* **52**, 727–733 (1984).
39. O'Donnell, M. E., Elias, P., Funnell, B. E. & Lehman, I. R. Interaction between the DNA-Polymerase and Single-Stranded DNA-Binding Protein (Infected Cell Protein-8) of Herpes-Simplex Virus-1. *J. Biol. Chem.* **262**, 4260–4266 (1987).
40. Makhov, A. M., Boehmer, P. E., Lehman, I. R. & Griffith, J. D. Visualization of the unwinding of long DNA chains by the herpes simplex virus type 1 UL9 protein and ICP8. *J. Mol. Biol.* **258**, 789–799 (1996).
41. Makhov, A. M., Taylor, D. W. & Griffith, J. D. Two-dimensional crystallization of herpes simplex virus type 1 single-stranded DNA-binding protein, ICP8, on a lipid monolayer. *Biochim. Biophys. Acta* **1701**, 101–108 (2004).
42. Makhov, A. M. et al. The bipolar filaments formed by herpes simplex virus type 1 SSB/recombination protein (ICP8) suggest a mechanism for DNA annealing. *J. Mol. Biol.* **386**, 273–279 (2009).
43. Mapelli, M., Panjekar, S. & Tucker, P. A. The crystal structure of the herpes simplex virus 1 ssDNA-binding protein suggests the structural basis for flexible, cooperative single-stranded DNA binding. *J. Biol. Chem.* **280**, 2990–2997 (2005).

44. Darwish, A. S., Grady, L. M., Bai, P. & Weller, S. K. ICP8 filament formation is essential for replication compartment formation during herpes simplex virus infection. *J. Virol.* **90**, 2561–2570 (2016).
45. Weerasooriya, S., DiScipio, K. A., Darwish, A. S., Bai, P. & Weller, S. K. Herpes simplex virus 1 ICP8 mutant lacking annealing activity is deficient for viral DNA replication. *Proc. Natl Acad. Sci. USA* **116**, 1033–1042 (2019).
46. Lopes, A., Amarir-Bouhram, J., Faure, G., Petit, M. A. & Guerois, R. Detection of novel recombinases in bacteriophage genomes unveils Rad52, Rad51 and Gp2.5 remote homologs. *Nucleic Acids Res.* **38**, 3952–3962 (2010).
47. Steczkiewicz, K., Prestel, E., Bidnenko, E. & Szczepankowska, A. K. Expanding diversity of firmicutes single-strand annealing proteins: a putative role of bacteriophage-host arms race. *Front. Microbiol.* **12**, 644622 (2021).
48. Kagawa, W. et al. Crystal structure of the homologous-pairing domain from the human Rad52 recombinase in the undecameric form. *Mol. Cell* **10**, 359–371 (2002).
49. Saotome, M. et al. Structural basis of homology-directed DNA repair mediated by RAD52. *iScience* **3**, 50–62 (2018).
50. Matsubara, K., Malay, A. D., Curtis, F. A., Sharples, G. J. & Heddle, J. G. Structural and Functional Characterization of the Red beta Recombinase from Bacteriophage lambda. *PLoS ONE* **8**, 78869 (2013).
51. Mythili, E., Kumar, K. A. & Muniyappa, K. Characterization of the DNA-binding domain of beta protein, a component of phage lambda red-pathway, by UV catalyzed cross-linking. *Gene* **182**, 81–87 (1996).
52. Wu, Z. et al. Domain structure and DNA binding regions of beta protein from bacteriophage lambda. *J. Biol. Chem.* **281**, 25205–25214 (2006).
53. Caldwell, B. J. et al. Crystal structure of the Redbeta C-terminal domain in complex with lambda Exonuclease reveals an unexpected homology with lambda Orf and an interaction with *Escherichia coli* single stranded DNA binding protein. *Nucleic Acids Res.* **47**, 1950–1963 (2019).
54. Jumper, J. et al. Highly accurate protein structure prediction with AlphaFold. *Nature* **596**, 583–589 (2021).
55. Punjani, A., Rubinstein, J. L., Fleet, D. J. & Brubaker, M. A. cryoSPARC: algorithms for rapid unsupervised cryo-EM structure determination. *Nat. Methods* **14**, 290–296 (2017).
56. Rosenthal, P. B. & Henderson, R. Optimal determination of particle orientation, absolute hand, and contrast loss in single-particle electron cryomicroscopy. *J. Mol. Biol.* **333**, 721–745 (2003).
57. Theobald, D. L., Mitton-Fry, R. M. & Wuttke, D. S. Nucleic acid recognition by OB-fold proteins. *Annu. Rev. Biophys. Biomol. Struct.* **32**, 115–133 (2003).
58. Holm, L. Using dali for protein structure comparison. *Methods Mol. Biol.* **2112**, 29–42 (2020).
59. Lee, C. H. & Charney, E. Solution conformation of DNA. *J. Mol. Biol.* **161**, 289–303 (1982).
60. Radding, C. M. & Shreffler, D. C. Regulation of lambda exonuclease. II. Joint regulation of exonuclease and a new lambda antigen. *J. Mol. Biol.* **18**, 251–261 (1966).
61. Illergard, K., Ardell, D. H. & Elovsson, A. Structure is three to ten times more conserved than sequence—a study of structural response in protein cores. *Proteins* **77**, 499–508 (2009).
62. Sugiman-Marangos, S. N., Weiss, Y. M. & Junop, M. S. Mechanism for accurate, protein-assisted DNA annealing by *Deinococcus radiodurans* DdrB. *Proc. Natl Acad. Sci. USA* **113**, 4308–4313 (2016).
63. Muniyappa, K. & Radding, C. M. The homologous recombination system of phage lambda. Pairing activities of beta protein. *J. Biol. Chem.* **261**, 7472–7478 (1986).
64. Zakharova, K., Caldwell, B. J., Ta, S., Wheat, C. T. & Bell, C. E. Mutational analysis of redbeta single strand annealing protein: roles of the 14 lysine residues in DNA binding and recombination in vivo. *Int. J. Mol. Sci.* **22**, 7758 (2021).
65. Armen, R., Alonso, D. O. & Daggett, V. The role of alpha-, 3(10)-, and pi-helix in helix→coil transitions. *Protein Sci.* **12**, 1145–1157 (2003).
66. Zhang, J., Pan, X. & Bell, C. E. Crystal structure of lambda exonuclease in complex with DNA and Ca(2+). *Biochemistry* **53**, 7415–7425 (2014).
67. Schindelin, J. et al. Fiji: an open-source platform for biological-image analysis. *Nat. Methods* **9**, 676–682 (2012).
68. Emsley, P., Lohkamp, B., Scott, W. G. & Cowtan, K. Features and development of Coot. *Acta Crystallogr D. Biol. Crystallogr* **66**, 486–501 (2010).
69. 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).
70. Williams, C. J. et al. MolProbity: More and better reference data for improved all-atom structure validation. *Protein Sci.* **27**, 293–315 (2018).
71. Liebschner, P. V. A. D. et al. Macromolecular structure determination using X-rays, neutrons and electrons: recent developments in Phenix. *Acta Crystallogr D. Biol. Crystallogr* **D75**, 861–887 (2019).
72. Suzek, B. E., Huang, H., McGarvey, P., Mazumder, R. & Wu, C. H. UniRef: comprehensive and non-redundant UniProt reference clusters. *Bioinformatics* **23**, 1282–1288 (2007).
73. Katoh, K. & Standley, D. M. MAFFT multiple sequence alignment software version 7: improvements in performance and usability. *Mol. Biol. Evol.* **30**, 772–780 (2013).
74. Waterhouse, A. M., Procter, J. B., Martin, D. M., Clamp, M. & Barton, G. J. Jalview Version 2—a multiple sequence alignment editor and analysis workbench. *Bioinformatics* **25**, 1189–1191 (2009).
75. Phillips, J. C. et al. Scalable molecular dynamics on CPU and GPU architectures with NAMD. *J. Chem. Phys.* **153**, 044130 (2020).
76. Jo, S., Kim, T., Iyer, V. G. & Im, W. CHARMM-GUI: a web-based graphical user interface for CHARMM. *J. Comput. Chem.* **29**, 1859–1865 (2008).
77. Lee, J. et al. CHARMM-GUI supports the Amber force fields. *J. Chem. Phys.* **153**, 035103 (2020).
78. Marti-Renom, M. A. et al. Comparative protein structure modeling of genes and genomes. *Annu. Rev. Biophys. Biomol. Struct.* **29**, 291–325 (2000).
79. Ivani, I. et al. Parmbsc1: a refined force field for DNA simulations. *Nat. Methods* **13**, 55–58 (2016).
80. Jorgensen, W. L., Chandrasekhar, J., Madura, J. D., Impey, R. W. & Klein, M. L. Comparison of simple potential functions for simulating liquid water. *J. Chem. Phys.* **79**, 926–935 (1983).
81. Brunger, A., Brooks, C. L. & Karplus, M. Stochastic boundary-conditions for molecular-dynamics simulations of St2 water. *Chem. Phys. Lett.* **105**, 495–500 (1984).
82. Feller, S. E., Zhang, Y. H., Pastor, R. W. & Brooks, B. R. Constant-pressure molecular-dynamics simulation - the langevin piston method. *J. Chem. Phys.* **103**, 4613–4621 (1995).
83. Andersen, H. C. Rattle - a velocity version of the shake algorithm for molecular-dynamics calculations. *J. Comput. Phys.* **52**, 24–34 (1983).
84. Darden, T., York, D. & Pedersen, L. Particle mesh Ewald - an N.Log(N) method for Ewald sums in large systems. *J. Chem. Phys.* **98**, 10089–10092 (1993).
85. Humphrey, W., Dalke, A. & Schulten, K. VMD: visual molecular dynamics. *J. Mol. Graph* **14**, 33–38 (1996).
86. Romo, T. D. & Grossfield, A. LOOS: an extensible platform for the structural analysis of simulations. *Annu. Int. Confer. IEEE Eng. Med Biol. Soc.* **2009**, 2332–2335 (2009).
87. Michaud-Agrawal, N., Denning, E. J., Woolf, T. B. & Beckstein, O. MDAnalysis: a toolkit for the analysis of molecular dynamics simulations. *J. Comput. Chem.* **32**, 2319–2327 (2011).

88. Mirdita, M. et al. ColabFold: making protein folding accessible to all. *Nat Methods* **19**, 679–682 (2022).

## Acknowledgements

We would like to thank Associate Professor Aaron Oakley for his input during atomic modeling, and Dr. Simon Brown at the Molecular Horizons Cryo-Electron Microscopy Facility and the MASSIVE (Multi-modal Australian ScienceS Imaging and Visualization Environment) support team for their technical contributions during data collection and processing, respectively. Dr. Tolun extends his thanks to Dr. Donald Court for his support during the early stages of this work. This work was supported by NHMRC Ideas grant (GNT1184012) and start-up funds provided by UOW to G.T.

## Author contributions

G.T. conceived and coordinated the study. G.T. and J.L.B. supervised the students T.P.N. and L.J.F. T.P.N., J.L.B., N.P.J., H.Y., and J.C.B. contributed to the experiments and data collection. T.P.N., G.T., N.P.J., H.Y., J.L.B. contributed to the analysis of the data and the interpretation of the results. G.T., J.L.B., T.P.N., N.P.J., H.Y., and L.J.F. wrote the original manuscript, which was revised by all authors.

## Competing interests

The authors declare no competing interests.

## Additional information

**Supplementary information** The online version contains supplementary material available at <https://doi.org/10.1038/s41467-022-33090-6>.

**Correspondence** and requests for materials should be addressed to Gökhan Tolun.

**Peer review information** *Nature Communications* thanks Wataru Kagawa, Kenan Murphy, Shenping Wu and the other, anonymous, reviewer(s) for their contribution to the peer review of this work.

**Reprints and permission information** is available at <http://www.nature.com/reprints>

**Publisher's note** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

**Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this license, visit <http://creativecommons.org/licenses/by/4.0/>.

© The Author(s) 2022