Engineering Microbiology 4 (2024) 100120

Contents lists available at ScienceDirect

Engineering Microbiology

journal homepage: www.elsevier.com/locate/engmic

Review

Half a century after their discovery: Structural insights into exonuclease and annealase proteins catalyzing recombineering

Lucy J. Fitschen[a,b], Timothy P. Newing[a,b], Nikolas P. Johnston[a,c], Charles E. Bell[d,*], Gökhan Tolun[a,b,*]

[a] School of Chemistry and Molecular Bioscience, and Molecular Horizons, University of Wollongong, Wollongong, NSW, Australia
[b] The ARC Training Centre for Cryo-electron Microscopy of Membrane Proteins, University of Wollongong, Wollongong, NSW, Australia
[c] Faculty of Science, University of Technology Sydney, Ultimo, NSW 2007, Australia
[d] Department of Biological Chemistry and Pharmacology, The Ohio State University College of Medicine, Columbus, OH 43210, United States

ARTICLE INFO

Keywords:
Recombineering
Redβ
λExo
RecE
RecT
Protein structure
Annealase
Exonuclease

ABSTRACT

Recombineering is an essential tool for molecular biologists, allowing for the facile and efficient manipulation of bacterial genomes directly in cells without the need for costly and laborious *in vitro* manipulations involving restriction enzymes. The main workhorses behind recombineering are bacteriophage proteins that promote the single-strand annealing (SSA) homologous recombination pathway to repair double-stranded DNA breaks. While there have been several reviews examining recombineering methods and applications, comparatively few have focused on the mechanisms of the proteins that are the key players in the SSA pathway: a 5′→3′ exonuclease and a single-strand annealing protein (SSAP or "annealase"). This review dives into the structures and functions of the two SSA recombination systems that were the first to be developed for recombineering in *E. coli:* the RecET system from *E. coli* Rac prophage and the λRed system from bacteriophage λ. By comparing the structures of the RecT and Redβ annealases, and the RecE and λExo exonucleases, we provide new insights into how the structures of these proteins dictate their function. Examining the sequence conservation of the λExo and RecE exonucleases gives more profound insights into their critical functional features. Ultimately, as recombineering accelerates and evolves in the laboratory, a better understanding of the mechanisms of the proteins behind this powerful technique will drive the development of improved and expanded capabilities in the future.

1. What is recombineering?

The ability to clone and edit genetic material is an essential component of life scientists' toolkit, allowing research in numerous fields, from molecular biology to biochemistry and cell biology to biophysics. One genome editing method known as recombineering, a portmanteau of **recombi**nation mediated genetic engi**neering**, allows DNA manipulation without restriction enzymes or other *in vitro* enzymatic treatments [1]. Recombineering was initially developed for editing DNA within *Escherichia coli* using bacteriophage proteins, taking advantage of either the bacteriophage lambda (phage λ) Red (identified via recombination-deficient mutations) recombination system [2–4] or the Rac prophage RecET system [5]. Both systems combine an exonuclease for resecting dsDNA ends in the 5′→3′ direction, with an annealase for binding the resulting 3′-ssDNA overhang and annealing it to a homologous ssDNA molecule. Since first demonstrated in *E. coli*, recombineering has been successfully implemented in many other bacteria (Table 1), often using the exonuclease and annealase proteins from a host-specific bacteriophage. Recombineering is also the basis for related techniques such as multiplex automated genomic engineering (MAGE) [6] that can rapidly evolve new bacterial strains with enhanced functions.

The Red and RecET phage systems have been exploited for recombineering due to their simple, streamlined, and highly efficient pathway for homologous DNA recombination known as single strand annealing (SSA). SSA is one of the three main pathways used in eukaryotic cells for the repair of dsDNA breaks, along with non-homologous end joining (NHEJ) and homologous recombination (HR) [39–43]. While numerous informative reviews covering different aspects of recombineering are currently available [1,44–46], most pay only limited attention to the structures and mechanisms of the proteins that are the key workhorses

Abbreviations and Nomenclature: AFM, Atomic Force Microscopy; CTD, C-terminal domain; EATR, Exonuclease-Annealase Two-component Recombinase = TCR (Two Component Recombinase); ERF, Essential Recombination Function; HSV-1, Herpes Simplex Virus 1; NS-EM, Negative Staining Electron Microscopy; NTD, N-terminal domain; SSA, Single-Strand Annealing; SSAP, Single-Strand Annealing Protein = Annealase; T2RE, Type 2 restriction endonuclease family of nucleases.

* Corresponding authors.
E-mail addresses: bell.489@osu.edu (C.E. Bell), gokhan_tolun@uow.edu.au (G. Tolun).

https://doi.org/10.1016/j.engmic.2023.100120
Received 5 July 2023; Received in revised form 15 September 2023; Accepted 19 September 2023
Available online 22 September 2023

Table 1
A list of the organisms in which recombineering has been reported. * Denotes a gene name rather than a protein. Multiple entries may exist for bacterial species where recombineering is reported with different EATR pairs. Similar tables have been created by others [1,7–9].

Bacterial Host Target	Exonuclease	Annealase	EATR Origin	Reference
Acinetobacter baumannii	ACINIS123_2462*	ACINIS123_2461*	*A. baumannii* strain IS-123	[10]
Agrobacterium tumefaciens	λExo	Redβ	Bacteriophage λ	[11]
Bacillus subtilis	N/A	GP35	phage SPP1	[12]
Burkholderia thailandensis	λExo	Redβ	Bacteriophage λ	[13]
Burkholderia pseudomallei	λExo	Redβ	Bacteriophage λ	[13]
Burkholderia sp. DSM 7029	Redα7029	Redβ7029	DSM 7029	[14]
Caulobacter crescentus	N/A	Redβ	Bacteriophage λ	[15]
Clostridium acetobutylicum	N/A	CPF0939*	*C. perfringens*	[16]
Collinsella stercoris	N/A	CspRecT	*C. stercoris* phage	[17]
Corynebacterium glutamicum	RecT	RecE	Rac Prophage	[18]
Corynebacterium glutamicum	OrfC	OrfB	*L. pneumophila*	[18]
Corynebacterium glutamicum	GP61	GP60	Phage Che9c of *M. smegmatis*	[18]
E. coli	N/A	CspRecT	*C. stercoris* phage	[17]
E. coli	λExo	Redβ	Bacteriophage λ	[4]
E. coli	RecE	RecT	Rac Prophage	[5]
Klebsiella pneumonia	N/A	CspRecT	*C. stercoris* phage	[17]
Lactobacillus brevis	RecE homolog	RecT homolog	*L. brevis* KB290	[19]
Lactobacillus casei	LCABL_13,060*	LCABL_13,040*	prophage PLE3	[20]
Lactobacillus plantarum	lp_0642*	lp_0640*	prophage P1	[21]
Lactobacillus reuteri	N/A	$RecT_1$	*L. reuteri*	[22]
Lactobacillus rhamnosus	N/A	LprRecT	*Lactobacillus reuteri* prophage	[15]
Lactococcus lactis	N/A	$RecT_1$	*L. reuteri*	[22]
Mycoplasma pneumoniae	N/A	GP35	phage SPP1	[23]
Mycobacterium smegmatis	Gp60	Gp61	Phage Che9c	[24]
Mycobacterium tuberculosis	Gp60	Gp61	Phage Che9c	[25]
Photorhabdus luminescence	Pluα	Pluβ	*P. luminescence*	[8]
Pseudomonas aeruginosa	N/A	PapRecT	*P. aeruginosa* phage	[17]
Pseudomonas aeruginosa	λExo	Redβ	Bacteriophage λ	[26]
Pseudomonas putida	N/A	Rec2	*P. putida*	[27]
Pseudomonas syringae	$RecE_{Psy}$	$RecT_{Psy}$	*P. syringae*	[28]
Saccharomyces cerevisiae	N/A	Redβ	Bacteriophage λ	[29]
Salmonella enterica	λExo	Redβ	Bacteriophage λ	[30]
Shigella sonnei	λExo	Redβ	Bacteriophage λ	[31]
Shigella flexneri	λExo	Redβ	Bacteriophage λ	[31]
Shigella dysenteriae	λExo	Redβ	Bacteriophage λ	[31]
Shewanella oneidensis	N/A	W3 Beta	*Shewanella* sp. W3–18–1	[32]
Sinorhizobium meliloti	λExo	Redβ	Bacteriophage λ	[33]
Staphylococcus aureus	N/A	EF2132*	*Enterococcus faecalis*	[34]
Vibrio natriegens	SXT-Exo	SXT-Beta	SXT mobile genetic element	[35]
Xenorhabdus stockiae	Pluα	Pluβ	*P. luminescence*	[8]
Xenorhabdus stockiae	XBJ1_1172*	XBJ1_1171*	N/A	[36]
Yersinia pseudotuberculosis	λExo	Redβ	Bacteriophage λ	[37]
Zymomonas mobilis	RecE	RecT	Rac Prophage	[38]

behind the method. As structural knowledge of a protein can dramatically improve our understanding of its function, this review will focus on the structures of the exonuclease and annealase proteins that have been determined to date, including the annealase structures reported during the past year. By digging deep into the structures of these proteins, we can understand not only how they function within their native bacterial hosts but also how we can continue to expand and improve recombineering in the future.

2. The roles of EATR proteins in single-strand annealing

Recombineering utilizes bacteriophage proteins that catalyze homologous DNA recombination. These proteins form an **E**xonuclease-**A**nnealase **T**wo-component **R**ecombinase system, or EATR. The term SynExo (**Syn**aptase **Exo**nuclease pair) has also been used, in which case synaptase is synonymous with annealase. The terms annealase, synaptase, recombinase, and single strand annealing protein (SSAP) have all been used to refer to the same group of proteins. Herein, we use the term annealase to describe proteins that bind to ssDNA and catalyze the annealing of two homologous ssDNA strands in an ATP-independent manner. Prominent examples include RecT from *E. coli* [47], Redβ from phage λ [48,49], Rad52 from yeast and humans [42], and ICP8 from Herpes Simplex Virus 1 (HSV1) [50,51]. Annealases typically do not catalyze DNA strand-invasion reactions (insertion of a ssDNA strand into a homologous dsDNA molecule) like the RecA and RAD51 recombinases that are ATP-dependent [52]. Instead, the two EATR proteins work in concert to catalyze DNA recombination by SSA: the exonuclease binds to a dsDNA end and caries out 5′→3′ end-resection to form a 3′-ssDNA overhang, to which the annealase binds and anneals it to a homologous ssDNA. The two steps of the reaction, end-resection and annealing, are coupled to each other via a protein-protein interaction between the exonuclease and annealase [53,54].

Before discussing the mechanisms of EATR proteins in recombineering, it is worth considering their natural roles in the propagation of the bacteriophage that typically encode them. In this regard, the Red system from phage λ has been studied in most detail. The Red genes are not required for viability of phage λ, but they significantly enhance (by 6 to 10-fold) the number of phage particles produced upon lysis [45]. Exactly how recombination promotes phage propagation is not fully understood, but roles in replication [55], generation of concatemeric genomes for viral genome packaging [56], repair of dsDNA breaks for CRISPR evasion [57], and generation of genetic diversity [58,59] have been proposed. Although the exact mechanisms by which EATR proteins promote recombination in cells are still under investigation, several models have been proposed.

Stahl et al. described two models for phage λ recombination in *E. coli,* one that is dependent on the host RecA DNA strand-exchange ATPase,

and another that is RecA-independent [60]. These models have been reviewed in excellent detail by Murphy [45]. Briefly, the RecA-dependent model, which is prevalent in non-replicating cells, starts with digestion of a dsDNA end by λExo, which loads Redβ onto the nascent 3′ ssDNA overhang [60]. The host RecFOR proteins then facilitate replacement of Redβ with RecA on the 3′-overhang, which promotes invasion of the 3′-overhang into a homologous dsDNA molecule, and recombination proceeds from there via the normal host double-strand break repair system [45,60].

The RecA-independent model, also known as SSA, is prevalent in replicating cells. In this model, λExo and Redβ again function in concert to form a 3′-overhang bound by Redβ, which in this case directly catalyzes annealing of the 3′-overhang to a homologous ssDNA molecule. This pathway requires two DNA breaks at non-allelic sites on separate phage λ chromosomes, which, upon end-resection, can produce 3′-overhangs with complementary regions that can be directly annealed to one another. Following annealing, the excess non-homologous overhanging strands are removed, and any gaps formed are filled in by a polymerase. Finally, a ligase can seal the remaining nick to result in a fully repaired functional dsDNA molecule [45,60].

A similar type of SSA pathway can ensue when a dsDNA break occurs between two directly repeated sequences on the DNA. In such a scenario, the two 3′-overhangs formed by end-resection will have homologous sequences that can be directly annealed to one another. This results in repair of the dsDNA break, but with deletion of one of the two repeats along with the sequence between them. Although the phage λ chromosome doesn't have directly repeated regions, a very similar type of SSA pathway can be highly significant in eukaryotic cells, where such repeats are common [61]. Moreover, homologs of the RecET and Red EATR proteins encoded on the IncC conjugative plasmid replicating in *Vibrio cholerae* have been demonstrated to use this type of SSA to evade dsDNA breaks formed by a host CRISPR system [57]. Thus, the EATR-promoted SSA pathway operating at repeated DNA sequences can be relevant to bacteriophage (or plasmid) propagation, and studies of the RecET and Red proteins have served as a model for understanding double-strand break repair by SSA in humans.

The recombination model that is most relevant to recombineering is the RecA-independent SSA model that relies on Redβ annealase [62,63]. Recombineering can either employ synthetic ssDNA oligonucleotides as the input DNA electroporated into cells (typically in the range of 35–100 nucleotides), or a dsDNA cassette that can be much longer (5 - 10 kb). The dsDNA cassette is typically generated by PCR using primers containing terminal homologies to the recombination target site. In the case of oligonucleotides, the exonuclease component of the EATR is not required for recombination: the annealase binds the input oligonucleotide directly and anneals it to the target site exposed as ssDNA at the lagging strand of a replication fork [64]. By contrast, recombineering with dsDNA as the input DNA requires end-resection by the exonuclease, and the resulting 3′-ssDNA overhang is bound by the annealase. Interestingly, it appears that in dsDNA recombineering, the exonuclease typically digests one complete strand of the input dsDNA, and Redβ anneals the intact opposing strand to the target site by the same mechanism as for short oligonucleotides, at the lagging strand of a replication fork [62,63]. While recombination via the classical SSA pathway (*i.e.* end-to-end annealing) can occur during recombineering, the efficiency is lower due to the requirement of appropriately positioned dsDNA ends. Hence, in most cases the annealing events during recombineering occur at the replication fork.

As mentioned above, the classical RecA-independent SSA pathway in phage λ described by Stahl et al. [60] was greatly stimulated in replicating cells, although the reason for this was unclear. It was suggested that the role of replication in these experiments could be to generate appropriately positioned dsDNA ends resulting from rolling circle replication [45,60]. However, this type of SSA model did not account for the observed level of recombination in a phage λ infection, and it is conceivable that some annealing events may occur at the replication fork, as seen for recombineering [45]. Models for phage λ recombination involving replication have thus been proposed, including one involving a Replisome Invasion/Template Switch [65]. It seems unlikely that the full 50 kb phage λ chromosome would be digested during productive recombination as described above for dsDNA recombineering, but λExo is highly processive and can digest full dsDNA substrates of that length *in vitro* [66,67]. In any case, the use of phage EATR proteins in recombineering, as well as structural and biophysical studies of them have clearly shed new light on the possible mechanisms of phage λ recombination that have been studied for so many years by geneticists [68].

Lastly, while EATR proteins form a complex with one another, they also interact with host proteins to facilitate recombination. Most prominently, Redβ from bacteriophage λ binds to *E. coli* single-stranded DNA binding protein (SSB) [15,69], which coats the ssDNA at the lagging strand of the replication fork to protect it from nucleases and control access of numerous replication proteins. This interaction with SSB is absolutely required for recombination *in vivo* [69], presumably to displace SSB and allow Redβ to gain access to the lagging strand. Redβ from phage λ also interacts with other proteins, including phage λ replication protein P, integrase, and antitermination protein [70], although the roles of these interactions are unclear. The interactions with host proteins are even more prevalent for EATR proteins from viruses that infect eukaryotic cells. For example, ICP8, the annealase from HSV-1, interacts with other proteins in the HSV-1 replisome, such as the UL9 origin-binding protein [71] and ICP27 that is essential for the regulation of viral gene expression [72].

In summary, structural information on EATR proteins and their interactions with other viral and host proteins is of broad interest for understanding multiple aspects of genome maintenance including replication, repair, and generation of genetic diversity. While the RecT and Redβ annealases that have been predominantly employed for recombineering have been studied for over half a century, the key structural insights into these proteins have come relatively recently, as summarized in the historical timeline in Fig. 1. The recent breakthroughs in annealase structures have put a new spotlight on these proteins and how understanding their structure can help unravel their function and lead to improvements in recombineering in the future.

2.1. Exonuclease structures

While recombineering with single-stranded oligonucleotides as the electroporated input DNA (often referred to as single-stranded oligonucleotide repair or ssOR) only requires the annealase [64], when the input DNA is double-stranded, both the annealase and exonuclease components of a specific EATR pair are needed. In addition to recombineering, exonucleases as stand-alone enzymes have been exploited for critical roles in other biotechnology applications such as generating ssDNA from dsDNA for PCR [86], CHIP-EXO protein-DNA foot-printing [87] and generating ssDNA for several biosensor applications (a few examples include [88–91]). Despite such various uses, little is known about how these proteins are evolutionarily related to each other, especially when compared to the work done for grouping annealases, discussed below [42,92].

The structures of the two main exonucleases used in *E. coli* recombineering, λExo from phage λ [69,78,82,83] and RecE from Rac prophage [80], have been determined by x-ray crystallography (Fig. 2). Remarkably, despite having limited sequence identity, both exonucleases form ring-shaped oligomers with central funnel-shaped channels, although λExo forms a trimer and RecE a tetramer. In both structures, the dsDNA is thought to enter at the wider end of the channel, such that the 5′-strand can feed into one of the active sites to be digested into mononucleotides. The 3′-overhang then exits out the back of the channel to tether the ring to the DNA as it moves forward digesting the 5′-strand [80,82,83]. This same oligomeric architecture has been seen for λExo, RecE, and for a third member of this family whose structure has

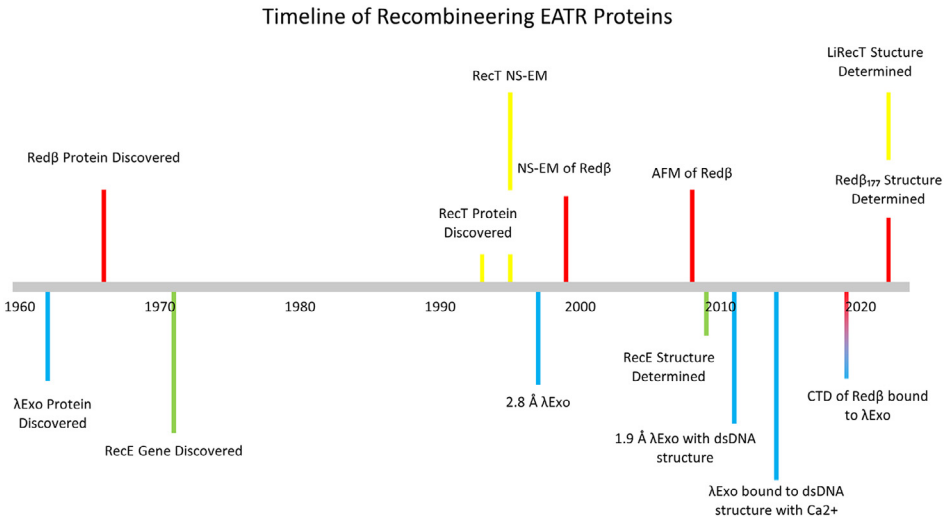

Fig. 1. Timeline of the discovery and structural advances of EATR proteins used in recombineering. References for each discovery are indicated in the figure [47,69,73–85]. Yellow indicates discoveries for RecT, red for Redβ, blue for λExo, and green for RecE.

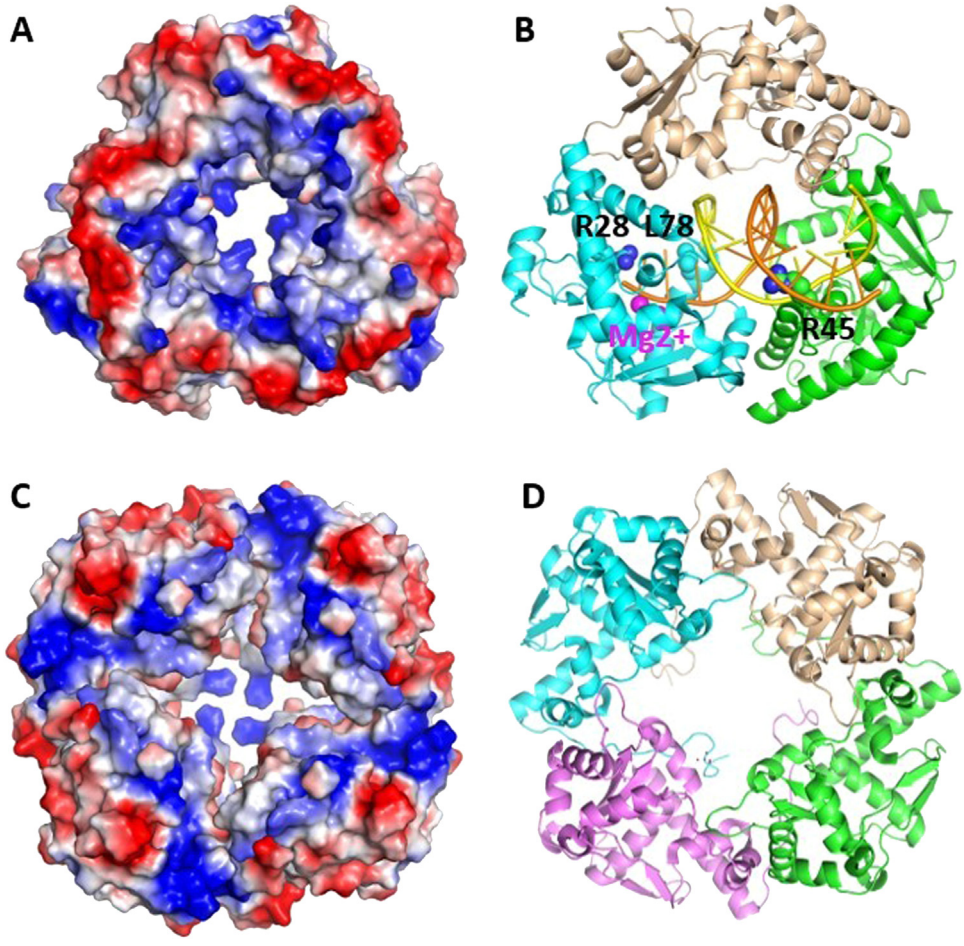

Fig. 2. Structures of λExo and RecE. A. Surface electrostatic view of the λExo trimer (PDBID:1AVQ) [78]. The wide end of the central channel to which the dsDNA substrate would enter is facing the viewer. B. Cartoon view of the λExo trimer complexed with dsDNA substrate (PDBID:3SM4) [82]. The 5′-end of the orange DNA strand binds in the active site of the cyan subunit, near two Mg^{2+} ions. Side chains discussed in the text are labeled. C, D. Surface electrostatics and cartoon view of the RecE nuclease domain tetramer (PDBID:3H4R) [80]. As for λExo in panels A and B, the view is looking down into the wide end of the central channel, to which the dsDNA would enter. Notice that the central channels in both structures narrow at the back end.

been determined, the alkaline Exonuclease from *Laribacter hongkongensis* [93,94].

2.2. *λExo structure*

λExo is a highly processive alkaline exonuclease that initiates digestion at dsDNA ends. The rate of dsDNA digestion is 5–40 nucleotides per second, as determined both at the single molecule level [66,67,95] and in bulk biochemical studies [82,96–98]. A peculiar feature of λExo is that it requires a 5′-phosphate on the dsDNA end for active digestion [75] yet binds to dsDNA with either 5′-OH or 5′-PO_4 ends with roughly equal affinity [99]. As the 5′-PO_4 is five covalent bonds removed from the phosphodiester bond that is cleaved in the reaction, its impact on catalytic activity but not on binding was perplexing. A clue as to the role of the 5′-phosphate came from mutagenesis studies indicating a pivotal role for Arg-28 in enzyme processivity, and an interaction of Arg-28 with the 5′-phosphate was suggested based on modeling [100].

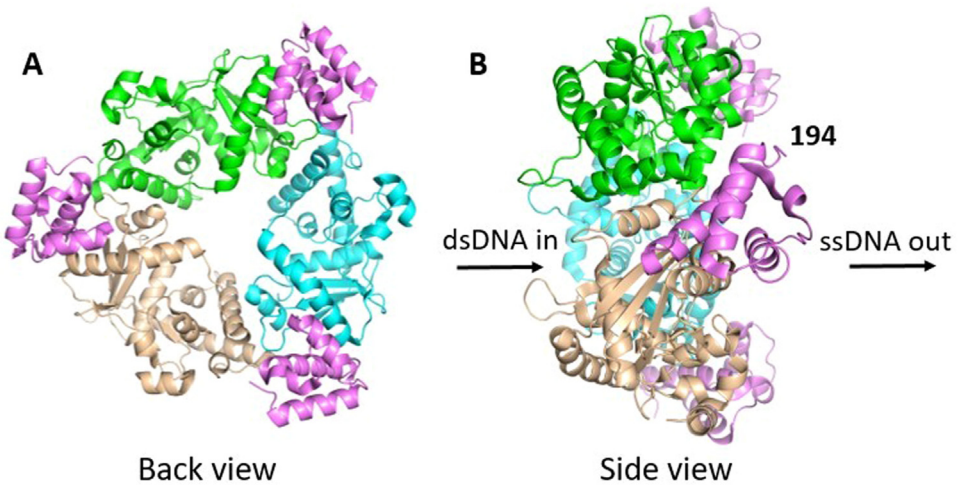

Fig. 3. Structure of λExo in complex with the Redβ CTD. A. view from the back of the λExo trimer (PDBID:6M9K) [69]. The dsDNA substrate would enter the ring from the back and the 3′-overhang ssDNA would come out towards the viewer. The three Redβ CTDs are colored magenta. B. Side view showing that the three Redβ NTDs and linkers (residues 1–193) would lie on the face of the λExo trimer from which the 3′-overhang ssDNA is extruded. The structure thus suggests a mechanism in which Redβ monomers are directly loaded onto the 3′-overhang as it is formed by λExo.

λExo was the first recombineering protein to have its crystal structure determined. Although first crystallized in 1985, the crystals at that time only diffracted to a 6 Å resolution [101]. It was not until 12 years later that a crystal structure was determined at 2.8 Å resolution, without DNA [78]. The structure revealed a ring-shaped homotrimer with a central channel of 30 Å at one end (Fig. 2a), enough to allow dsDNA to enter, but only 15 Å at the other end, allowing only ssDNA to exit [78]. The proposed DNA binding mode nicely explained the high processivity of λExo, as the ring-shaped trimer would be topologically linked to the DNA molecule as it moves along digesting it.

Over a decade later, the structure of λExo in complex with DNA substrate was determined [82] (Fig. 2b). The crystallized complex contained a 12-bp duplex with a 5′-phosphorylated 2-nt overhang at one end (a 14-mer/12-mer), the inactive K131A variant of λExo to prevent DNA digestion, and the Mg^{2+} ions that are required for nuclease activity. The structure showed that the DNA is indeed bound to the central channel, but significantly tilted to place the end of the DNA with the 2-nt overhang into one of the three active sites. The two nucleotides at the 5′ end of the DNA are bent away from the duplex and inserted into an active site cleft, while the 3′-OH of the opposing strand is positioned to exit out the back of the trimer.

The unwinding of the DNA is mediated by apolar residues, including Leu-78 that wedge into the base pairs to separate them. The 5′-phosphate of the DNA is indeed bound at the end of the active site to Arg-28, while the scissile bond is bound to two Mg^{2+} ions held in place by crucial acidic active site residues. The structure, which visualizes the nucleophilic water molecule that is poised for attack [82], supports a classic two-metal nuclease mechanism [102] characteristic of the type 2 restriction endonuclease (T2RE) family [103], also known as the PD-(D/E)XK family [104]. Three loops of λExo, one from each subunit, extend from the rim of the central channel to contact the downstream portion of the dsDNA substrate. The Arg-45 side chain from one of the three loops inserts into the minor groove of the DNA and is proposed to help the enzyme keep on track. In support of this role, mutation of Arg-45 to Ala almost completely disrupts cleavage activity [82,105].

Based on this structure, an “electrostatic ratchet” model for processive digestion was proposed in which the interaction of the 5′-phosphate on the DNA with Arg-28 at the end of the active site is key to moving the enzyme forward. As each mononucleotide is cleaved from the 5′-end and released with Mg^{2+} out the rear portal on the trimer, the newly generated 5′-phosphate of the next nucleotide on the DNA would be attracted to the positively charged pocket containing Arg-28. The hydrophobic wedge formed by Leu-78 is proposed to help unwind the base pairs as the enzyme moves along the DNA, and the Arg-45 side chain is thought to act as a rudder to help the trimer track along the minor groove of the downstream portion of the DNA.

The most recent crystal structure of λExo shows a trimer bound to three copies of the Redβ CTD, resolved to 2.3 Å [69] (Fig. 3). This structure provided the first direct insights into the architecture of the λExo-Redβ EATR complex and is remarkably consistent with a model in which the role of the interaction is to load the Redβ annealase directly onto the 3′-overhang that is formed by λExo during digestion [45,106]. Further details of this interaction will be examined below in Section 2.5.

2.3. RecE structure

While λExo is a 226 amino acid protein, RecE is a much larger 866 amino acid protein that contains a C-terminal nuclease domain (residues 564–866) and an N-terminal domain of unknown function. The nuclease domain can substitute genetically for the full-length protein [107], although full-length RecE has enhanced activity for recombineering involving linear-linear (end-to-end) SSA recombination *in vivo* [108]. The crystal structure of the RecE nuclease domain was determined at 2.8 Å in the absence of DNA in 2009 [80]. The RecE fold has a core topology similar to λExo and a common set of conserved active site residues. Intriguingly, the RecE monomers pack into the tetramer in essentially opposite orientations as λExo, relative to the end of the channel at which the DNA would enter. This suggests that although RecE and λExo are evolutionarily related at the tertiary structure level, their similar quaternary structures (RecE tetramer and λExo trimer) likely evolved independently from a monomeric ancestor. Clearly, a ring-shaped structure with a tapered central channel is a fundamental architectural feature for this processive 5′→3′ exonuclease enzyme family.

Each subunit of RecE has a channel that contains an active site connecting with a positively charged portal that could allow for the release of mononucleotides as they are cleaved (Fig. 2c). The structure was determined without DNA in the presence of Ca^{2+}, which supports DNA binding but not cleavage. Although only one Ca^{2+} ion is bound per active site, two Mg^{2+} ions are presumably needed for cleavage. The set of critical active site residues is primarily conserved between RecE and λExo, with one notable exception: Glu-85 of λExo is replaced by His-652 in RecE. This residue is also histidine in the C-terminal nuclease domain of RecB of the *E. coli* RecBCD complex, another member of the T2RE family. The role of this residue in catalysis is not yet clear, but it could help to stabilize the 3′-OH leaving the group after hydrolysis.

Another difference between RecE and λExo is that RecE contains much longer loops projecting out from the rim of the central channel, presumably to capture the dsDNA substrate. These loops, formed by residues 665–698 of RecE, are largely disordered in the crystal structure and are not part of the final refined model. One of our laboratories

(Bell) successfully crystallized the RecE nuclease domain in complexes with different lengths of DNA. However, the DNA could never be visualized, presumably because it did not sit down in a unique orientation relative to the crystal packing interactions. The DNA-binding loops did, however, become partially visualized in these structures (data not shown).

In summary, the RecE and λExo structures show several common features that appear fundamental to the processive nuclease activity required for 5′→3′ end-resection. These features also appear to be conserved for the additional structures of related exonuclease proteins of the phage recombination systems that have been determined.

2.4. Evolutionary analysis of recombineering exonucleases by sequence alignments

To gain further insights into exonuclease function, we analyzed sequence conservation in both λExo and RecE using the 2000 hit blast search results against the UniProt ref90 database. Following multiple sequence alignment and quality control of these datasets, the final MSAs consisted of 1347 sequences for λExo and 183 for RecE. Many of the RecE sequences were eliminated once sequences with greater than 90% similarity were clustered. Both λExo and RecE belong to the PD-(D/E)XK phosphodiesterase superfamily, a highly diverse group of proteins with homologs present in all domains of life [103,104]. The superfamily consists primarily of nucleases, including processive exonucleases such as λExo, RecE, *E. coli* RecB, and the herpesvirus alkaline nuclease UL12, as well as many restriction endonucleases including those used in traditional cloning techniques. As with most members of this superfamily, λExo and RecE have a conserved core fold consisting of a four-stranded, mixed β-sheet flanked by α-helices, with α-β-β-β-α-β topology [104]. Embedded within this fold are the conserved aspartate, glutamate, and lysine residues that give the PD-(D/E)XK family its name.

In our analysis, the active site residues of both proteins appear to be highly conserved, and in most cases identical across all constituent sequences. In contrast, other regions of the protein are more variable (Fig. 4a & b, Supplementary Figures 1 & 2). Notably, λExo and RecE differ in the composition of their active site residues. λExo displays a highly conserved PD-EXK active site structure (Fig. 4c), whereas RecE has the alternate structure PD-DXK. In both alignments, conserved positively charged residues flank the active site and are thought to facilitate binding to the DNA substrate.

Of particular interest is the great disparity in length between the λExo and RecE families. While the PD-(D/E)XK-like domain spans the entire length of the λExo sequences, it comprises only the C-terminal segment of RecE. A review of the hits retrieved by RecE reveals a heterogeneity in sequence length, with some showing homology across the entire length of RecE including its large N-terminal domain. In contrast, others like λExo consist of only a single exonuclease domain (Supplementary Figure 1). As expected, the C-terminal PD-(D/E)XK domain showed the greatest conservation across all alignment regions prior to truncation, possibly explaining the higher number of sequences eliminated during the clustering step.

A review of the hits retrieved by BLAST search of λExo identified many sequences identified as homologs of YqaJ, a domain from a known EATR pair found in the skin element of *Bacillus subtilis* [109]. Other homologs of YqaJ include Chu exonuclease of the *B. subtilis* phage SPP1, which forms an EATR pair with its partner annealase GP35 [12,109].

Also of note, the elements that encode λExo and RecE have different reproductive methods. λExo is encoded within a phage that can undergo lytic reproduction, whereas RecE is present in a defective prophage replicating with the host. This difference in reproduction method could have a marked effect on sequence evolution. Phage λ has a generation time of ~7.7 phages min^{-1} [110], which is >20x faster than that of *E. coli,* which is estimated at ~0.3 bacterium min^{-1} [111]. The more rapid evolution of λExo could account for the higher level of sequence similarity for the RecE family in our analysis. Alternatively, the differences in similarity could reflect the limited number of RecE sequences available in the current UniProt Database.

2.5. The lambda phage EATR complex: λExo+Redβ

The interaction between the two phage EATR proteins has been known for over half a century. In fact, the Redβ protein was discovered during the purification of λExo, as the two proteins were seen to co-purify with an apparent 1:1 stoichiometry [53]. While the nuclease activity of λExo had been well known [74,75], the function of Redβ was not established until nearly a decade later when it was discovered that Redβ could promote the annealing of homologous ssDNA strands [48]. Although Redβ can function independently, somewhat higher annealing activity was observed in the presence of λExo [48]. The reason for this is still not apparent. There is actually a third protein in the Red system known as γ-protein (also referred to as Gam), encoded by the *gam* gene [112–114]. The γ-protein is often not present in genomes encoding a typical EATR pair, including the *E. coli* Rac prophage encoding RecE and RecT, and it appears to take on a more supplementary role. In phage λ, the *gam* gene is required for transitioning from the early to the late stage of viral infection [115], but the γ-protein does not appear to interact with any other phage λ recombination proteins [116]. Instead, it binds to the host RecBCD helicase/exonuclease complex to prevent it from digesting dsDNA ends [112–114], which are present on the linear form of the λ phage genome [112]. The γ-protein can efficiently inhibit both nuclease activities of RecBCD, including its exonuclease activity on dsDNA and its ssDNA and endonuclease activity on ssDNA [113,114]. Crystal structures of γ-protein reveal a small alpha-helical dimer, and a cryo-EM structure of γ-protein in complex with RecBCD has been determined [117,118]. Red-mediated recombination with linear dsDNA can occur without γ-protein *in vivo* [7]. Still, γ-protein is typically included in recombineering strains with active RecBCD to prevent the destruction of linear duplex DNA.

While there is currently no structure of a complete EATR complex, there have been attempts to model what the λ phage EATR complex could look like (Fig. 5). One of the first models, proposed by Tolun in 2007, considered the available biochemical and stoichiometric data [106] and was comprised of four λExo trimers bound to a dodecameric ring of Redβ (Fig. 5a). This complex would presumably load onto a dsDNA end through one of the λExo trimers. According to the model, as the 5′-strand is digested, the exposed ssDNA would be bound to the N-terminal domains of the associated Redβ subunits [45,106]. Although the purified EATR complex has a 1:1 stoichiometry [53], a higher concentration of Redβ compared to λExo would be required, as λExo would only need to be stoichiometric with dsDNA ends, whereas a higher concentration of Redβ would be required to saturate the long ssDNA overhangs generated. Presumably, Redβ monomers detach from the EATR complex to form a large oligomeric nucleoprotein complex by coating the nascent 3′-overhang ssDNA, primed for annealing for carrying out SSA [81]. Indeed, expression of Redβ at higher levels than λExo leads to a significant improvement of recombination efficiency, whereas an excess of λExo over Redβ decreases recombination levels [54]. A similar relationship was observed for RecE and RecT, suggesting that the two EATR pairs work by similar mechanisms [54]. Despite this trend, when Redβ and λExo are expressed for recombineering from their natural P_L promoter on a pSIM5 vector, Western blot analysis with a polyclonal antibody raised to both proteins indicates that similar levels of each are expressed [69,119]. Whether this reflects their levels at the end of a phage λ infection when recombination is most active is uncertain.

While the first EATR model was primarily based on biochemical data, Newing et al. [81] proposed a model based on newly available structural data (Fig. 5b). This model incorporated the structures of λExo bound to DNA [82] and to the CTD of Redβ [69], combined with the $Red\beta_{177}$ cryo-EM structure [81]. AlphaFold 2 was used to predict a structure of the linker region of Redβ (residues 178–193), which has not yet been

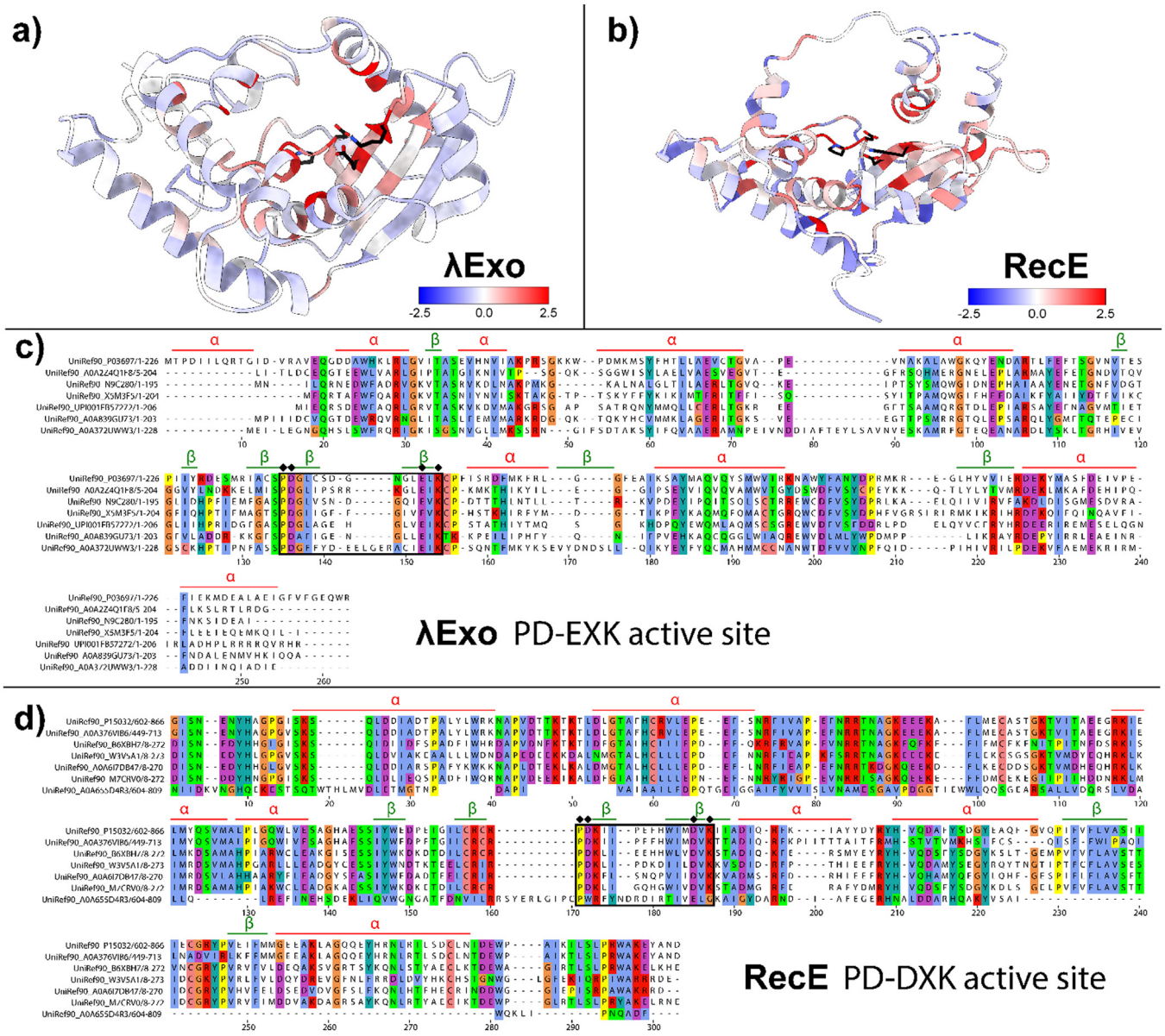

Fig. 4. Sequence conservation of *λ*Exo and RecE: Sequence conservation of a) *λ*Exo (PDBID:1AVQ) 1347 sequences, and b) RecE (PDBID:3H4R) 183 sequences [78,80]. Scale values represent the entropy-based AL2CO conservation index, where blue and red correspond to the lowest and highest sequence conservation, respectively. Active site residues are indicated in black stick representation. Representative MSA is shown for c) *λ*Exo and d) RecE that summarize the full 2000 hit analysis. Secondary structural features, mapped from atomic structures of the reference sequences, are indicated above the MSA, with *α*-helices in orange and *β*-sheets in green. Active site residues are indicated above the MSA by a black diamond, and the active site region is bound by a black box. The consensus active site for each family is provided at the end of each respective MSA.

resolved experimentally. This model also assumed higher levels of Red*β* than *λ*Exo as the long 3′-ssDNA overhang generated by *λ*Exo trimer digestion would require multiple Red*β* monomers to form a continuous protein-DNA filament as seen in the sample used for determining the structure of Redβ_{177}.

Here we propose a third possible model for the phage *λ* EATR complex, generated using AlphaFold 2 (Fig. 5c,d). The model contains three *λ*Exo and three Red*β* subunits, and retains the signature 1:1 ratio, with *λ*Exo forming the characteristic trimer, three Red*β* N-terminal domains interacting with each other as in the Redβ_{177} structure, and three Red*β* C-terminal domains bound to *λ*Exo as in the crystal structure of the complex [69]. While this model is asymmetric in how the N-terminal domains are positioned, the Red*β* linker region is likely to be flexible enough to allow for the conformational differences. Interestingly, the cleft on the Red*β* N-terminal domain that contains DNA in the cryo-EM structure is occupied by a new *α*-helix from the linker region that is generated by AlphaFold 2 [81]. This *α*-helix (labelled *α*-7 and colored pink in Fig. 2b) would block DNA binding and could conceivably control how Red*β* monomers assemble on the nascent ssDNA that is generated by *λ*Exo, acting as a regulatory element. While there is still no experimentally determined structure of the full phage *λ* EATR complex, the insights we gain from each new structure and model help us in assembling the pieces.

2.6. Annealase proteins

There are many distinct types of proteins with annealase activity found in nature, as first mapped out by Iyer et al. who proposed three distinct superfamilies grouped around ERF (essential recombination function), RecT/Red*β*, and Rad52. Each family was predicted to

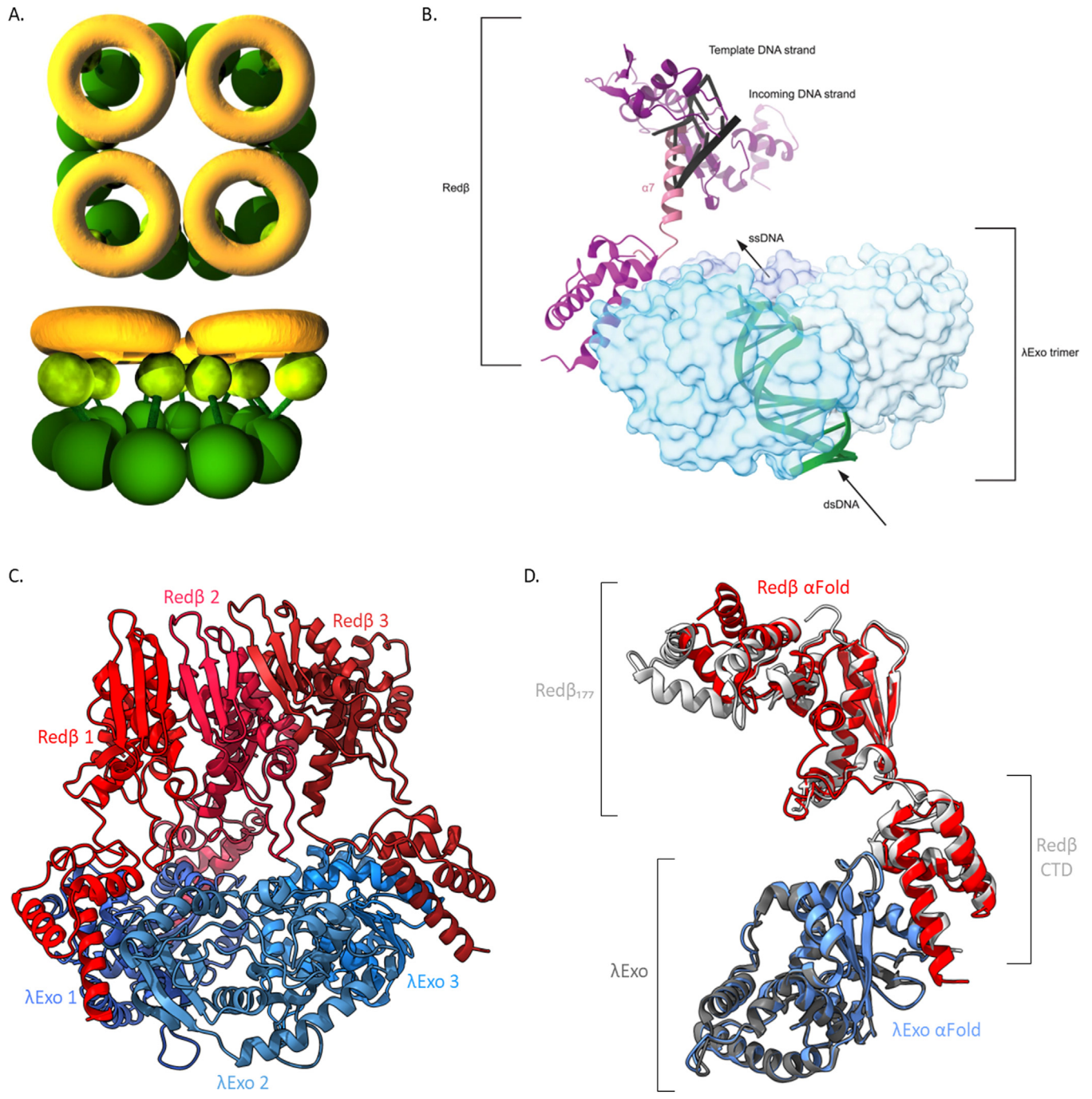

Fig. 5. Proposed models of the EATR complex of Redβ and λExo. (A) is an early model, reproduced with author permission from Tolun [106], showing four λExo trimers (yellow) bound to the C-terminal domains of Redβ (lime) that form a larger ring with their N-terminal domains (darker green). (B) is a reproduction with author permission of the model proposed by Newing et al. [81] showing a composite structural model of the complex, with Redβ highlighted in purple and λExo in blue. The α-7 helix shown in pink is not part of the experimental structure, but was generated by AlphaFold 2. It would block the DNA-binding groove and may somehow serve in a regulatory role. AlphaFold 2 was used to predict the structure of Redβ and λExo (C), showing a λExo trimer (blue) bound to three Redβ monomers (red). (D) highlights the interaction of single λExo and Redβ monomers with the experimentally determined structures of λExo (PDBID:6M9K) in dark gray, and of the C-terminal domain of Redβ (PDBID:6M9K) and the first 177 amino acids of Redβ (PDBID:7UJL) in light gray [69,81].

have a different core fold and a distinct pattern of sequence conservation [42]. Lopes et al. later proposed a different grouping based on Rad52-like, Gp2.5-like, and Rad51-like sequences [92]. Most recently, seven annealase families were proposed including Sak3, Sak4, Rad52/22, ERF, RecT/Redβ, Gp2.5, and RecA [120]. The latter two groupings included Rad51/RecA family proteins that have annealase activity, but primarily function in ATP-dependent DNA strand invasion and exchange for homologous recombination [52]. Similarly, Gp2.5 is a single-stranded DNA binding protein from bacteriophage T7 that presumably has annealase activity as a side effect of ssDNA-binding [121]. Due in part to the diversity of annealase proteins, we have yet to arrive at a consensus mechanism for how they catalyze DNA annealing. Based on their distinct core folds and presumably different evolutionary origins, the different types of annealase proteins could indeed operate by different mechanisms. Of the 7 families described most recently, only three have

representative high-resolution structures available, namely the Rad52, RecA/RAD51, and RecT/Redβ families.

Recombineering has primarily been developed and optimized for use in *E. coli*, and the RecET and phage λ Red proteins have evolved to function in *E. coli.* The annealase activity is now known to depend on an interaction with the host SSB protein [15,69], which will vary in sequence in different hosts. Therefore, it can be challenging to predict if λ Red or RecET will be functional in a given bacterial species of interest. However, recombineering can be expanded to new organisms by mining for EATR proteins from a bacteriophage (or prophage) that infect them (Table 1). Moreover, Redβ from phage λ functions efficiently as an annealase for recombineering in close relatives of *E. coli* including *Salmonella enterica* [30]. The interaction between annealases and host SSB proteins largely involves the last ~9 residues of SSB [15,69], which is the site for interaction of numerous *E. coli* host replication proteins [122]. Altering this sequence has allowed the portability of a given annealase into a new bacterium of interest [15], providing a simple means to increase the efficiency of recombineering in new bacterial hosts. Knowledge of how RecT and Redβ operate can also benefit our understanding of the annealase mechanism, for which there has been a general lack of structural information, particularly for the relevant protein-DNA complexes. While structures with ssDNA substrates have been available for eukaryotic annealases including Rad52 and ICP8, the recent structures of RecT and Redβ [81,84] have been determined in complex with a duplex annealing intermediate, and therefore provide important new insights into the annealing mechanism, as described below.

2.7. Structure of Redβ

While Redβ was discovered over half a century ago [74], structural investigations only began only ~20 years ago. The first structures reported by Passy et al. used negative staining electron microscopy (NS-EM) [79] and revealed oligomeric rings in the absence of DNA and larger rings with ssDNA. Left-handed helical filaments were observed when Redβ was mixed with heat-denatured double stranded DNA, which was the first indication of a structural transition upon annealing [79]. Almost ten years after Passy et al.'s findings, further investigations using atomic force microscopy (AFM) [85] revealed similar helical filaments in the presence of two complementary ssDNA sequences, but disperse monomers bound to a single ssDNA sequence. A model for annealing was proposed in which a clamped dimer of Redβ stabilizes a nucleus of complementarity from which annealing can propagate [85].

While the work of Erler et al. confirmed that Redβ forms ring-like structures in the absence of DNA, it more clearly showed structures resembling a split-lock washer, with a gap or a slight overlap between monomers at one end of each ring [85]. Unlike NS-EM, which gives 2D projections, AFM imaging is sensitive to height, which explains how split-lock washers were detected. As there were some differences in the oligomeric complexes seen by different low resolution imaging methods, a high-resolution structure of Redβ was clearly needed to resolve the discrepancies.

The first atomic structure of Redβ was of its C-terminal domain, determined in complex with a λExo trimer. The overall architecture of the complex supported a model in which Redβ is loaded onto ssDNA during DNA end-resection by λExo [69]. From a mutational analysis of the λExo-CTD interface, a second role for the CTD in binding to the host SSB protein was discovered. The two interactions were found to use an overlapping site and are thus likely to be mutually exclusive. A 'hand-off' model was proposed in which the interaction with λExo loads Redβ onto the first ssDNA (the 3′-overhang formed by λExo), while the subsequent interaction with SSB localizes the initial Redβ-ssDNA complex to the lagging strand of the replication fork, where it can scan the lagging strand for a sequence that is complementary to the first ssDNA [64,69]. The structure of the Redβ CTD is also significant because the other available structure of Redβ that would become available only includes its N-terminal DNA-binding domain [81].

Most recently, cryo-EM revealed the structure of the Redβ N-terminal domain (NTD) that is responsible for DNA binding and oligomerization (Fig. 6b). Rather dramatically, the structure captured a helical filament of Redβ in complex with a novel intermediate of DNA annealing that has an unusual conformation of duplex DNA [81]. The structure used a truncated form of Redβ that only included its first 177 amino acids (out of 261 in native Redβ). The cryo-EM 2D class averages showed 1- and 2- start helical filaments, with the start of a helix denoting the number of threads that are found per turn of a helix [81]. While 2-

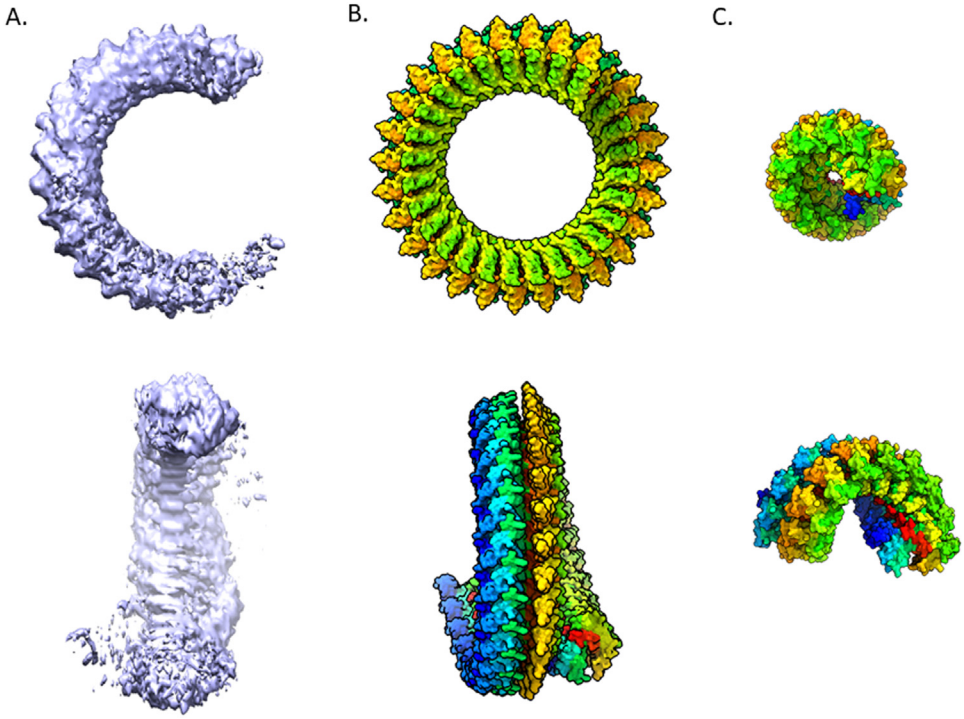

Fig. 6. Comparison of Redβ and LiRecT cryo-EM structures. (A) shows a preliminary cryo-EM map of full-length Redβ, showing the incomplete ring shape. This is compared to the Redβ_{177} structure (B), which has been assembled into the 'split-lock washer' conformation (PDBID:7UJL) [81]. A comparison of these two structures highlights the extra density in (A), which matches up with the overlapping parts of the helix in (B). LiRecT (C) structure is also shown, with both a top and side view (PDBID:7UB2) [84].

start filaments have also been observed for the ICP8 annealase from HSV-1, their functional role is not clear [123,124]. The 1-start filaments of Redβ_{177} on the other hand suggested a compelling mechanism for annealing.

The Redβ_{177} structure demonstrates the molecular mechanism of how this protein, and most likely its homologs in the Redβ/RecT family, anneal DNA. The cryo-EM sample was created by sequentially adding complementary 27 nt ssDNA oligonucleotides, a common approach for forming annealed DNA intermediates [49,79,85,119]. For the first time, a protein was visualized in complex with a conformation of DNA with a ribbon-like planar structure, as opposed to the typical B-form double helix [81]. While this conformation was first seen for Redβ, it was quickly confirmed for RecT, lending strong support for functional significance and a conserved mechanism of annealing [84]. The structure clearly revealed a site size of 4 base pairs of DNA for each Redβ monomer. However, each Redβ monomer contacts nucleotides from 6 consecutive base pairs of the bound DNA [81].

This structure also revealed the mechanism of oligomerization of Redβ subunits. While it had been well known that Redβ forms oligomeric complexes [79,85] the Redβ_{177} structure showed that oligomerization is stabilized mainly through electrostatic interactions between monomers, with patches of positive and negative charges that mirror each other on opposite faces [81]. It was also noted that some of the residues involved in oligomerization help to stabilize the DNA-binding pocket, reinforcing the idea that oligomerization is inherently coupled to DNA-binding [81]. Caldwell et al. [119] examined the oligomeric structures formed by Redβ with different DNA substrates. They found that while the binding of Redβ to ssDNA gave rise to a wide distribution of different species, the sequential addition of two complementary oligonucleotides formed a much more distinct and stable complex with Redβ. While the relationship between oligomerization and DNA binding is still not fully understood, the new structural data support the conclusion that they are coupled.

Although not resolved in the cryo-EM map, residues 130–139 of Redβ form a so-called “finger loop.” Molecular Dynamics simulations suggested that this loop is highly flexible and moves between open and closed conformations to control access of the DNA strands to the binding groove on Redβ. By holding the second (incoming) ssDNA strand within the groove to sample the more deeply bound template ssDNA, while allowing it to disengage if the match is not perfect, this loop could play a crucial role in decreasing the sequence homology search from 3D (*i.e.*, diffusing into and out of the DNA-binding site from the solution) to 1D (*i.e.*, sliding in the DNA-binding site, along the template ssDNA strand). Further experiments will be needed to test this model.

2.8. The structure of RecT from a prophage of Listeria innocua

For many years, one of our laboratories (Bell) had been attempting to crystallize RecT/Redβ homologs (and fragments thereof) alone and in complex with DNA. Although several crystals with DNA were obtained, none diffracted x-rays well enough to permit structure determination. With the advent of methods for high-resolution cryo-EM, we prepared grids of full-length Redβ alone and in complex with different lengths of ssDNA and sequentially added complementary DNAs. These complexes invariably gave spiral C-shaped structures that sat in the plane of the ice (Fig. 6a). Although the 3D structures of the complexes could not be determined due to strong preferred orientation, wedge-shaped subunits were readily apparent. Curiously, the structures contained many more subunits (>20) than expected from earlier ring-shaped complexes of λ Redβ seen by negative stain EM, with previous EM studies estimating between 11 and 18 subunits per ring [79,85]. In hindsight, these appear very similar to the cryo-EM structure solved at high resolution for the Redβ_{177} NTD, in terms of the approximate number of subunits per turn, overall diameter, and shallow helical pitch [81]. This suggests that the complexes formed by the NTD as seen by cryo-EM [81] are likely to be very similar to the complexes formed by the full-length Redβ. However, since the complexes of full-length Redβ did not stack end to end as for the NTD, they did not form the longer filaments in the plane of the ice to be amenable to helical reconstruction.

To approach this problem, we took advantage of several full-length RecT/Redβ homologs we had purified over the years to screen for crystallization. While three of these gave similar C-shaped structures as λ Redβ, one of them, RecT from the A118 prophage of *Listeria innocua (LiRecT)*, gave distinct helical filaments when mixed with two complementary 83-mer oligonucleotides that were added to the protein sequentially [84]. The filaments appeared strikingly similar to those of full-length Redβ when mixed with long heat-denatured dsDNA as seen by negative staining EM [79]. Single particle analysis of these filaments resulted in the 3.4 Å 3D reconstruction of a single 1.5-turn filament containing 18 LiRecT subunits and 83 bp of duplex DNA (Fig. 6c). The DNA was bound in a highly extended, unwound, and flattened conformation to a narrow, positively charged groove that runs along the outer surface of the filament. Although the 2D class averages and 3D reconstruction converged on a single 1.5-turn filament with 10 subunits per turn, considerably longer filaments were observed in the raw cryo-EM images, suggesting that the filaments can stack end to end. While these longer filaments were not used in the single particle analysis, they would likely be amenable to helical reconstruction, although this has not yet been attempted.

The LiRecT monomer shares a common core fold with Rad52 and Redβ. LiRecT contains three inserted segments: an N-terminal helical bundle that forms inter-subunit interactions at the upper rim of the filament, a β_A-β_B hairpin inserted between β_3 and β_4 that forms inter-subunit interactions at the lower rim of the filament, and a pair of helices (αD-αE) inserted after β_5 that forms the lower rim of the DNA binding groove. Only residues 34 to 244 of each LiRecT monomer were visible in the reconstruction. The C-terminal residues (245–271) could form a small domain analogous to the CTD of Redβ for interacting with its exonuclease partner and host SSB protein. However, at only 26 residues (instead of ~70), it would be considerably smaller.

While the LiRecT mixed with two complementary 83-mer DNA strands formed a stable complex that was solved to 3.4 Å resolution, a lower resolution structure of LiRecT was determined in a complex with a single 83-mer oligonucleotide. This complex also forms left-handed helical filaments, although the filaments were less well-organized and did not stack end-to-end. Although the bound DNA could not be resolved clearly enough to model, there was a considerable amount of density in the groove. One clear and informative feature of this complex was that the N-terminal lobes of each LiRecT subunit were apparently disordered, as their corresponding density was weak. This observation was intriguing as it suggested a mechanism in which the N-terminal lobes clamp down on the duplex only when a fully complementary strand of ssDNA is incorporated into the complex.

Native mass spectrometry of LiRecT-DNA complexes [84] largely confirmed the numbers of subunits in the LiRecT-ssDNA and LiRecT-dsDNA complexes seen by cryo-EM. Interestingly, a mixture of two different types of LiRecT-ssDNA complexes was observed by native MS: some with one copy of 83-mer ssDNA and 10–12 LiRecT subunits, and some with two copies of the same 83-mer ssDNA and 17–18 LiRecT subunits. While the functional significance of these different types of complexes is yet to be established, we hypothesize that the smaller complexes with a single copy of ssDNA represent initial LiRecT-ssDNA complexes, while the larger complexes represent attempts at annealing at sites of partial complementarity along the 83-mer. In agreement with this hypothesis, the larger complexes are similar in composition to the helical filaments formed with two-complementary DNAs and were likely the species resolved by single-particle analysis of the complex formed when only one 83-mer ssDNA was added to LiRecT.

Based on the LiRecT structures, a model for annealing was proposed in which a cluster of ~7 LiRecT subunits (one for each 5 nt of ssDNA) forms on ssDNA with the ssDNA bound to the inner site. This initial LiRecT-ssDNA complex can sample additional strands of ssDNA and

bring them into the groove to attempt to pair with the initially bound template ssDNA. When a complementary match is found, the two DNAs can form a closely paired (but flattened) duplex, and the N-terminal lobe of LiRecT can clamp down on the duplex to stabilize and consolidate annealing. This forms an additional and extensive amount of inter-subunit interactions at the N-terminal lobes (above the DNA binding groove), to form the highly stable complex seen by gel shift and single-molecule experiments [49,85,125].

2.9. Comparison of Redβ and RecT

The recent boom in annealase structures creates an opportunity to discover common features that may be important for function. While both Redβ and RecT have been studied since their discovery in the 1960s and 1990s, respectively, it was only in 2022 that their structures became available [81,84]. Despite their high structural similarity, RecT and Redβ have only ~10% sequence identity, and the phage annealase family is extremely divergent [92]. While RecT and Redβ are the defining members of the RecT/Redβ annealase superfamily [42] and were expected to have similar structures, visualization of their common core fold has cemented this relationship. This β-β-α-β-β-β-α fold is seen in both LiRecT and Redβ_{177}, with the most conserved portion being the β-β-β-α core comprised of an α-helix crossing diagonally over three antiparallel β-strands. This β-β-β-α section of the fold was first observed in the structure of the N-terminal domain (NTD) of Rad52 [126,127], which solidifies the notion that this entire group of annealases is evolutionarily related from bacteriophages to humans [92].

A notable structural divergence in phage annealase proteins is in their C-terminal domains (CTD), which are believed to bind the exonuclease partner and host SSB protein. The first interaction was seen in a crystal structure of the CTD of Redβ bound to λExo [69]. Whether other phage annealases including RecT use a similar CTD to bind to their respective partner proteins is unknown. The interaction has been biochemically and functionally confirmed for the *E. coli* RecT/RecE EATR, with RecT binding to RecE but not to λExo, demonstrating specificity [54]. While there have been attempts to model what a complete EATR complex containing λExo, Redβ, and DNA could look like [81], exactly how this complex forms and its mechanistic details will remain unknown until its structure is determined. Such a structure would be expected to explain how the full-length proteins assemble with a 1:1 stoichiometry, despite λExo forming trimers and Redβ forming rings, split-lock washers, or continuous helical filaments.

The bound DNA conformation is another important similarity between the Redβ and RecT structures. Both complexes have the DNA duplex held in a highly extended, flattened, and completely unwound conformation. However, there are some notable differences. First, each LiRecT monomer binds to 5 bp of DNA duplex, whereas Redβ (and RAD52) bind to 4 bp (or nt) per monomer. Second, the Redβ filaments contain 27 subunits (108 nt) per helical turn, whereas the LiRecT filaments are much more tightly wound with 10 subunits (50 bp) per helical turn. Both have a 5-nt (or 4-nt) repeating pattern of extended DNA, where the protein wedges into the DNA every 5 (or 4) nt to form an extended gap where the base pairs are completely un-stacked.

The details of the protein-DNA and protein-protein interactions that build up the filaments of LiRecT and Redβ differ greatly, due to the extremely low sequence similarity. Nonetheless, a striking similarity is that the inner template DNA strand, added first to each protein to form the annealing intermediate, forms considerably more extensive interactions with the protein than the outer (incoming) DNA strand that was added second. The outer (incoming) strand is mainly held in place through normal Watson-Crick base pairs with the inner (template) strand. In the case of Redβ_{177} there could be additional interactions involving the disordered finger loop that was not visualized in the structure. Indeed, the finger loop harbors three basic residues (K132, R134, R137) that could potentially interact with the incoming strand, to supplement its Watson-Crick base pair interactions with the template strand.

2.10. Recombineering annealases and human Rad52

While Redβ and RecT are essential to *E. coli* recombineering with both dsDNA and ssDNA, there are a multitude of other annealases found in nature from organisms other than bacteriophages. The idea of recombineering in eukaryotes is not new, and by characterizing the annealases from other organisms, there may be new insights into how this can be achieved. Unfortunately, these proteins are poorly understood, and many annealases such as Erf and Sak4, do not have available structures. Despite the dearth of phage annealase structures, several structures are available for human Rad52 [126–129]. Al-Fatlawi et al. [130] compared the structures of Rad52, Redβ_{177}, and LiRecT, and described a common annealase fold as the β-β-β-α core. When Al-Fatlawi et al. compared LiRecT and Redβ_{177} to the crystal structure of Rad52, the conserved region of phage annealases was expanded to encompass a β-β-α-β-β-β-α fold as the DNA binding region [130]. Not only was this fold visible in both structures, but it had previously been predicted to be conserved among annealases [92]. This β-β-α-β-β-β-α fold was also seen within the LiRecT structure, again showing homology between all the current annealase structures. Fig. 7 demonstrates the conserved nature of this fold, highlighting the distinctive secondary structures in LiRecT, Redβ_{177}, and Rad52. These findings demonstrate a structural link between eukaryotic and phage annealases, and show that they bind to ssDNA substrate, and possibly duplex intermediate, via similar mechanisms.

The available crystal structures of the NTD of Rad52, including a structure with ssDNA substrate, reveal planar undecamermeric rings [126–129]. By contrast, negative stain EM of full length Rad52 showed heptameric rings [131]. Helical filaments have never been observed for Rad52, despite the functional (and now structural) similarities with Redβ and LiRecT. We envision two possibilities to account for this. First, as proposed by Al-Fatlawi et al., the model for Rad52 involving ring-to-ring annealing *in trans* may need to be re-examined [130]. It is conceivable that Rad52 could form filaments under conditions yet to be identified, such as at lower concentrations that are likely to be more relevant *in vivo*. Along these lines, Rad52 has only been visualized at high resolution when bound to ssDNA substrate, while RecT and Redβ have only been visualized bound to duplex intermediate. RecT and Redβ could indeed function as oligomeric rings at earlier stages of their annealing reactions, such as when bound to an ssDNA substrate. While there is no high-resolution structure of Redβ or RecT in the ring form, AFM [85] and NS-EM [79] have shown that Redβ can form rings or split-lock washers. The extensive range of oligomeric complexes observed for Redβ (summarized in Table 2) indicates a dynamic oligomerization yet to be completely unraveled. It is clear that the relationship between the annealase fold and the multimeric complexes formed by annealases still has much to be explored.

2.11. Recombineering annealases and other DNA binding proteins

While Rad52 is one of the best-studied eukaryotic annealases, there are many more annealases found throughout nature, including from viruses like HSV1 that infect humans. The ICP8 annealase from HSV-1 has been well characterized, and its crystal structure was determined as a monomer [132]. Although the fold of ICP8 is completely unrelated to Rad52 and the RecT/Redβ bacteriophage annealases, ICP8 functions with an exonuclease partner called UL12 to form an analogous EATR complex. ICP8 can stimulate gene targeting in human cell lines [133,134], but the efficiency was less than 0.5%. Thus, further optimization will be needed for eukaryotic recombineering to become more useful in practice.

The most noticeable difference between ICP8 and the phage annealases is the larger size of the ICP8 monomer (128 kDa) compared to

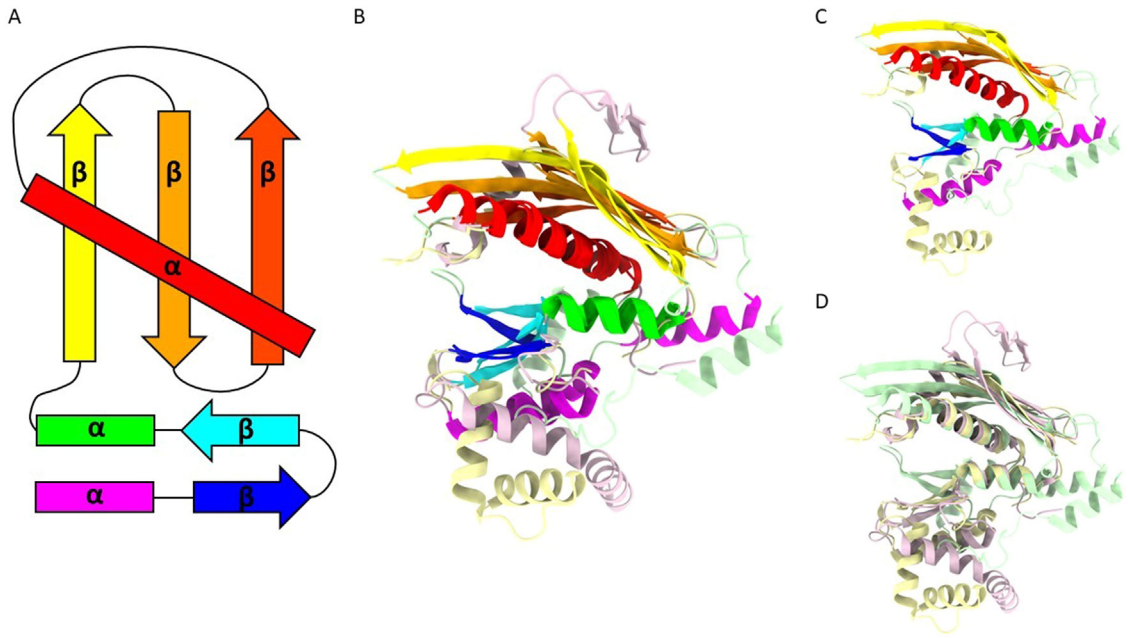

Fig. 7. Visualization of the comparison between Redβ_{177}, LiRecT, and Rad52 structures. (A) shows a topographical representation of the α-β-β-α-β-β-α fold, inspired by [130]. Each of the secondary structures is highlighted and corresponds to the appropriate structure in (B), which superimposes Redβ_{177} (yellow, PDBID:7UJL), LiRecT (pink, PDBID:7UB2), and Rad52 (green, PDBID:8H1P)[82,85,130]. These structures were also represented without the highlighted secondary structures (D) and without Rad52 (C).

Table 2
The number of subunits in oligomeric complexes formed by Redβ observed using different techniques.

Number of subunits	Technique Used	DNA Substrate	Reference
15 - 16	TEM	30 nt oligonucleotide	[79]
18	TEM	Heat-denatured calf thymus DNA	[79]
11 – 12	TEM	No DNA	[79]
14 ± 3 per turn	AFM	83 nt, 123 nt, and 163 nt oligonucleotides[1]	[85]
11 - 12	AFM	No DNA	[85]
14	SEC-MALS	38 nt ssDNA	[119]
21.1	SEC-MALS	83 nt ssDNA	[119]
11.3	SEC-MALS	38 nt oligonucleotides[1]	[119]
18.2	SEC-MALS	83 nt oligonucleotdies[1]	[119]
7.7 - 9.1	SEC-MALS	No DNA	[119]
11.5	AUC	38 nt oligonucleotides[1]	[119]
12.5	AUC	38 nt ssDNA	[119]
9 - 13	AUC	No DNA	[119]
11	Native-MS	38 nt oligonucleotides[1]	[119]
22	Native-MS	83 nt oligonucleotides[1]	[119]
9	Native-MS	38 nt ssDNA	[119]
11	Native-MS	83 nt ssDNA	[119]
≈ 12	Native-MS	No DNA	[119]

[1] Two complementary oligonucleotides added to Redβ sequentially to form complex with annealed duplex.

RecT (29.7 kDa) and Redβ (29.7 kDa). This could reflect the involvement of ICP8 in additional cellular processes including replication [135] and SSB-like ssDNA binding [136]. While RecT and Redβ form left-handed helical filaments upon DNA binding and annealing [81,84], ICP8 can form filaments in the absence of DNA [51,124]. The shape of the ICP8 filament is also different. While RecT and Redβ form predominantly single, one-start helical filaments, ICP8 forms anti-parallel double (bipolar) filaments. The DNA binding groove on ICP8 is located along the central helical axis, between the two anti-parallel filaments. This suggests a *trans* mechanism of SSA, and contrasts with the *cis* SSA mechanism inferred from the LiRecT and Redβ structures, in which the two paired DNAs are bound together within the same (outer) groove on a single filament. The inter-subunit packing in ICP8 is also different from the phage annealases (and from Rad52), reflecting its different protein fold and evolutionary origin. All of these differences indicate a larger structure, and more elaborate oligomerization behavior for ICP8. Further details of ICP8 will be presented in an upcoming article from the Tolun group.

2.12. Future directions for recombineering and annealases

Knowledge of how the phage λ and Rac prophage EATR protein pairs operate is essential for understanding the SSA pathway and improving recombineering for future biotechnology applications. Recent work on the phage EATR proteins has now provided at least one structure for Redβ, λExo, RecT, and RecE. While this work has highlighted key similarities among these proteins, many questions remain. How do the exonuclease and annealase partners come together to form a complete EATR machine? During SSA, does the EATR function as a stable, intact complex, or do the exonuclease and annealase components associate and dissociate dynamically during SSA? While the exonucleases are needed at a stoichiometric ratio to the DNA ends available, annealases are presumably needed at higher levels to bind the length of the nascent ssDNA generated by the exonucleolytic digestion. Yet, the EATR complexes appear to have a 1:1 stoichiometry. How does this binding ratio impact the SSA reaction? While we now have structures for each of the four main individual EATR proteins, no structures are available for a full EATR complex. There have been attempts to model what this complex could look like [81], but only limited structural data to support them [69]. Therefore, one of the next frontiers will be to determine the structures of complete EATR complexes bound to relevant DNA substrates and annealing intermediates that reflect the different stages of the SSA reaction.

An exciting goal is to identify EATRs from viruses infecting organisms from different kingdoms of life. This would facilitate the development and expansion of recombineering to new organisms. While hundreds of exonucleases and annealases have been identified by various bioinformatics approaches, most appear to function only in bacteria, and they are typically very distantly related to one another at the sequence level. Outside of the four main proteins discussed here, few other EATR

proteins have been extensively characterized *in vitro*. While eukaryotic recombineering is an exciting goal for the future, further structural characterization of the EATRs of HSV1 and related viruses will be needed to bring it to fruition.

3. Conclusion

This review summarized the literature on EATRs catalyzing SSA used in recombineering from a structural biology perspective, and provides new results on the conservation of EATRs. The recent annealase structures determined by our groups revealed striking molecular mechanistic details and evolutionary connections. These structures can be exploited to improve recombineering efficiency in the future. Yet, the many questions listed above remain. Future structures of EATR complexes bound to their DNA substrates will unravel further mechanistic details about how SSA takes place during recombineering. The potential of developing methods for recombineering in eukaryotic cells using the EATRs of HSV1 and related viruses is an enticing goal for the future. Given the lack of a requirement for a specific type of DNA sequence, recombineering by SSA could potentially be as impactful as CRISPR, if not more, although further development is needed.

4. Methods

4.1. Comparison of sequence conservation between λExo and RecE

A multiple sequence alignment (MSA) was prepared based on 2000 hit BLASTsearch against the Uniref90 protein database. The subsequent hits were then initially aligned using the FFT-NS-2 algorithm in MAFFT v7.52, and the alignments truncated to the region of the query sequence containing the PD-(D/E)XK endonuclease-like domain, as identified in InterPro (1–226 for Exo, 602–866 for RecE) [137,138]. Truncated sequences were clustered by 90% global sequence identity using CD-HIT [139]. Further quality control was completed, including removing sequences less than 150 aa and removing outliers based on a PCA within Jalview v2.11 [140]. The final MSA were re-aligned using the L-INS-i algorithm in MAFFT v7.52 [138]. The resultant MSA were then used to project sequence conservation onto the protein structures in ChimeraX v1.61 using the entropy-based AL2CO method [141,142]. To visualize the sequence variation, secondary structure location, and identity of the active site residues, a small subsection of the original MSA was created containing the query sequence and 6 additional sequences representing the variation retrieved from the full 2000 hit analysis. MSA figures were generated using the pyMSAviz package (available from: https://moshi4.github.io/pyMSAviz/). Structural predictions were performed using the ColabFold [143] implementation of AlphaFold2 multimer version 3.

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

CRediT authorship contribution statement

LJF, CEB, and GT wrote the manuscript. TPN and NPJ performed the sequence analysis on RecE and λExo exonuclease families, prepared Fig.4, and wrote the text for Section 2.4 and Section 4.1. CEB prepared Fig. 2 and Fig. 3. LJF prepared all other figures and tables.

Acknowledgments

This work was funded by the National Science Foundation Grant MCB-2212951 (to CEB) and NHMRC Ideas grant APP1184012/GNT1184012 (to GT).

Supplementary Materials

Supplementary material associated with this article can be found, in the online version, at doi:10.1016/j.engmic.2023.100120.

References

[1] T.M. Wannier, P.N. Ciaccia, A.D. Ellington, G.T. Filsinger, F.J. Isaacs, K. Javanmardi, M.A. Jones, A.M. Kunjapur, A. Nyerges, C. Pal, M.G. Schubert, G.M. Church, Recombineering and MAGE, Nat. Rev. Methods Primers 1 (2021), doi:10.1038/s43586-020-00006-x.
[2] K.C. Murphy, Use of bacteriophage lambda recombination functions to promote gene replacement in Escherichia coli, J. Bacteriol. 180 (1998) 2063–2071, doi:10.1128/JB.180.8.2063-2071.1998.
[3] K.A. Datsenko, B.L. Wanner, One-step inactivation of chromosomal genes in Escherichia coli K-12 using PCR products, Proc. Natl. Acad. Sci. 97 (2000) 6640–6645, doi:10.1073/PNAS.120163297.
[4] D. Yu, H.M. Ellis, E.C. Lee, N.A. Jenkins, N.G. Copeland, D.L. Court, An efficient recombination system for chromosome engineering in Escherichia coli, Proc. Natl. Acad. Sci. U.S.A. 97 (2000) 5978–5983, doi:10.1073/pnas.100127597.
[5] Y. Zhang, F. Buchholz, J.P.P. Muyrers, A.Francis Stewart, A new logic for DNA engineering using recombination in Escherichia coli, Nat. Genet. 20 (1998) 123–128, doi:10.1038/2417.
[6] H.H. Wang, F.J. Isaacs, P.A. Carr, Z.Z. Sun, G. Xu, C.R. Forest, G.M. Church, Programming cells by multiplex genome engineering and accelerated evolution, Nature 460 (2009) 894–898, doi:10.1038/nature08187.
[7] S. Datta, N. Costantino, X. Zhou, D.L. Court, Identification and analysis of recombineering functions from Gram-negative and Gram-positive bacteria and their phages, Proc. Natl. Acad. Sci. U.S.A. 105 (2008) 1626–1631, doi:10.1073/pnas.0709089105.
[8] J. Yin, H. Zhu, L. Xia, X. Ding, T. Hoffmann, M. Hoffmann, X. Bian, R. Müller, J. Fu, A.F. Stewart, Y. Zhang, A new recombineering system for Photorhabdus and Xenorhabdus, Nucleic Acids Res. 43 (2015) e36, doi:10.1093/nar/gku1336.
[9] B.J. Caldwell, C.E. Bell, Structure and mechanism of the Red recombination system of bacteriophage λ, Prog. Biophys. Mol. Biol. 147 (2019) 33–46, doi:10.1016/j.pbiomolbio.2019.03.005.
[10] A.T. Tucker, E.M. Nowicki, J.M. Boll, G.A. Knauf, N.C. Burdis, M.Stephen Trent, B.W. Davies, Defining gene-phenotype relationships in acinetobacter baumannii through one-step chromosomal gene inactivation, mBio 5 (2014) 1–9, doi:10.1128/MBIO.01313-14.
[11] S. Hu, J. Fu, F. Huang, X. Ding, A.F. Stewart, L. Xia, Y. Zhang, Genome engineering of Agrobacterium tumefaciens using the lambda Red recombination system, Appl. Microbiol. Biotechnol. 98 (2014) 2165–2172, doi:10.1007/s00253-013-5412-x.
[12] Z. Sun, A. Deng, T. Hu, J. Wu, Q. Sun, H. Bai, G. Zhang, T. Wen, A high-efficiency recombineering system with PCR-based ssDNA in Bacillus subtilis mediated by the native phage recombinase GP35, Appl. Microbiol. Biotechnol. 99 (2015) 5151–5162, doi:10.1007/s00253-015-6485-5.
[13] Y. Kang, M.H. Norris, B.A. Wilcox, A. Tuanyok, P.S. Keim, T.T. Hoang, Knockout and pullout recombineering for naturally transformable Burkholderia thailandensis and Burkholderia pseudomallei, Nat. Protoc. (2011) 1085–1104 2011 6:8. 6, doi:10.1038/nprot.2011.346.
[14] X. Wang, H. Zhou, H. Chen, X. Jing, W. Zheng, R. Li, T. Sun, J. Liu, J. Fu, L. Huo, Y. zhong Li, Y. Shen, X. Ding, R. Müller, X. Bian, Y. Zhang, Discovery of recombinases enables genome mining of cryptic biosynthetic gene clusters in Burkholderiales species, Proc. Natl. Acad. Sci. U.S.A. 115 (2018) E4255–E4263, doi:10.1073/PNAS.1720941115.
[15] G.T. Filsinger, T.M. Wannier, F.B. Pedersen, I.D. Lutz, J. Zhang, D.A. Stork, A. Debnath, K. Gozzi, H. Kuchwara, V. Volf, S. Wang, X. Rios, C.J. Gregg, M.J. Lajoie, S.L. Shipman, J. Aach, M.T. Laub, G.M. Church, Characterizing the portability of phage-encoded homologous recombination proteins, Nat. Chem. Biol. 17 (2021) 394–402, doi:10.1038/s41589-020-00710-5.
[16] H. Dong, W. Tao, F. Gong, Y. Li, Y. Zhang, A functional recT gene for recombineering of Clostridium, J. Biotechnol. 173 (2014) 65–67, doi:10.1016/j.jbiotec.2013.12.011.
[17] T.M. Wannier, A. Nyerges, H.M. Kuchwara, M. Czikkely, D. Balogh, G.T. Filsinger, N.C. Borders, C.J. Gregg, M.J. Lajoie, X. Rios, C. Pál, G.M. Church, Improved bacterial recombineering by parallelized protein discovery, Proc. Natl. Acad. Sci. U.S.A. 117 (2020) 13689–13698, doi:10.1073/pnas.2001588117.
[18] Y. Chang, Q. Wang, T. Su, Q. Qi, Identification of phage recombinase function unit in genus Corynebacterium, Appl. Microbiol. Biotechnol. 105 (2021) 5067–5075, doi:10.1007/s00253-021-11384-x.
[19] H. Huang, X. Song, S. Yang, Development of a RecE/T-Assisted CRISPR–Cas9 Toolbox for Lactobacillus, Biotechnol. J. 14 (2019) 1800690, doi:10.1002/biot.201800690.
[20] Y. Xin, T. Guo, Y. Mu, J. Kong, Identification and functional analysis of potential prophage-derived recombinases for genome editing in Lactobacillus casei, FEMS Microbiol. Lett. 364 (2017) 243, doi:10.1093/femsle/fnx243.
[21] P. Yang, J. Wang, Q. Qi, Prophage recombinases-mediated genome engineering in Lactobacillus plantarum, Microb. Cell Fact. 14 (2015) 1–11, doi:10.1186/s12934-015-0344-z.
[22] J.P. Van Pijkeren, R.A. Britton, High efficiency recombineering in lactic acid bacteria, Nucleic Acids Res. 40 (2012) e76, doi:10.1093/nar/gks147.
[23] C. Piñero-Lambea, E. Garcia-Ramallo, S. Martinez, J. Delgado, L. Serrano, M. Lluch-Senar, Mycoplasma pneumoniae Genome Editing Based on Oligo Recombineer-

ing and Cas9-Mediated Counterselection, ACS Synth. Biol. 9 (2020) 1693–1704, doi:10.1021/acssynbio.0C00022.
[24] J.C. Van Kessel, G.F. Hatfull, Efficient point mutagenesis in mycobacteria using single-stranded DNA recombineering: Characterization of antimycobacterial drug targets, Mol. Microbiol. 67 (2008) 1094–1107, doi:10.1111/J.1365-2958.2008.06109.X.
[25] J.C. van Kessel, G.F. Hatfull, Recombineering in Mycobacterium tuberculosis, Nat. Methods 4 (2007) 147–152, doi:10.1038/nmeth996.
[26] B.-2199-9-20. pdf Lesic L.G. Rahme, Use of the lambda Red recombinase system to rapidly generate mutants in Pseudomonas aeruginosa, BMC Mol. Biol. 9 (2008) 1–9, doi:10.1186/1471-2199-9-20.
[27] T. Aparicio, A. Nyerges, I. Nagy, C. Pal, E. Martínez-García, V. de Lorenzo, Mismatch repair hierarchy of Pseudomonas putida revealed by mutagenic ss-DNA recombineering of the pyrF gene, Environ. Microbiol. 22 (2020) 45–58, doi:10.1111/1462-2920.14814.
[28] B. Swingle, Z. Bao, E. Markel, A. Chambers, S. Cartinhour, Recombineering Using RecTE from Pseudomonas syringae, Appl. Environ. Microb. 76 (2010) 4960–4968, doi:10.1128/AEM.00911-10.
[29] E.M. Barbieri, P. Muir, B.O. Akhuetie-Oni, C.M. Yellman, F.J. Isaacs, Precise Editing at DNA Replication Forks Enables Multiplex Genome Engineering in Eukaryotes, Cell 171 (2017) 1453–1467, doi:10.1016/j.cell.2017.10.034.
[30] K. Bunny, J. Liu, J. Roth, Phenotypes of lexA Mutations in Salmonella enterica: Evidence for a Lethal lexA Null Phenotype Due to the Fels-2 Prophage, J. Bacteriol. 184 (2002) 6235–6249, doi:10.1128/JB.184.22.6235-6249.2002.
[31] R.T. Ranallo, S. Barnoy, S. Thakkar, T. Urick, M.M. Venkatesan, Developing live Shigella vaccines using λ Red recombineering, FEMS Immunol. Med. Microbiol. 47 (2006) 462–469, doi:10.1111/j.1574-695X.2006.00118.x.
[32] A.D. Corts, L.C. Thomason, R.T. Gill, J.A. Gralnick, A new recombineering system for precise genome-editing in Shewanella oneidensis strain MR-1 using single-stranded oligonucleotides, Sci. Rep. 9 (2019) 1–10, doi:10.1038/s41598-018-37025-4.
[33] J. Yang, Q. Zhang, G. Zhang, G. Shang, Recombineering-Mediated Sinorhizobium meliloti Rm1021 Gene Deletion, Curr. Microbiol. 80 (2023) 76, doi:10.1007/s00284-023-03188-1.
[34] K. Penewit, E.A. Holmes, K. McLean, M. Ren, A. Waalkes, S.J. Salipante, Efficient and scalable precision genome editing in Staphylococcus aureus through conditional recombineering and CRISPR/Cas9-mediated counterselection, mBio 9 (2018) e00067-18, doi:10.1128/mBio.00067-18.
[35] H.H. Lee, N. Ostrov, M.A. Gold, G.M. Church, Recombineering in Vibrio natriegens, Biorxiv (2017), doi:10.1101/130088.
[36] X. Huang, Y. Sun, S. Liu, Y. Li, C. Li, Y. Sun, X. Ding, L. Xia, Y. Hu, S. Hu, Recombineering using RecET-like recombinases from Xenorhabdus and its application in mining of natural products, Appl. Microbiol. Biotechnol. 106 (2022) 7857–7866, doi:10.1007/s00253-022-12258-6.
[37] A. Derbise, B. Lesic, D. Dacheux, J.M. Ghigo, E. Carniel, A rapid and simple method for inactivating chromosomal genes in Yersinia, FEMS Immunol. Med. Microbiol. 38 (2003) 113–116, doi:10.1016/S0928-8244(03)00181-0.
[38] Y. Wu, T. Li, Q. Cao, X. Li, Y. Zhang, X. Tan, RecET recombination system driving chromosomal target gene replacement in Zymomonas mobilis, Electron. J. Biotechnol. 30 (2017) 118–124, doi:10.1016/j.ejbt.2017.10.005.
[39] R. Ceccaldi, B. Rondinelli, A.D. D'Andrea, Repair Pathway Choices and Consequences at the Double-Strand Break, Trends Cell Biol. 26 (2016) 52–64, doi:10.1016/J.TCB.2015.07.009.
[40] R. Bhargava, D.O. Onyango, J.M. Stark, Regulation of Single-Strand Annealing and its Role in Genome Maintenance, Trends Genet. 32 (2016) 566–575, doi:10.1016/J.TIG.2016.06.007.
[41] E. Mladenov, S. Magin, A. Soni, G. Iliakis, DNA double-strand-break repair in higher eukaryotes and its role in genomic instability and cancer: Cell cycle and proliferation-dependent regulation, Semin. Cancer Biol. 37–38 (2016) 51–64, doi:10.1016/J.SEMCANCER.2016.03.003.
[42] L.M. Iyer, E.V. Koonin, L. Aravind, Classification and evolutionary history of the single-strand annealing proteins, BMC Genomics [Electronic Resource] 3 (2002), doi:10.1186/1471-2164-3-8.
[43] S. Ramakrishnan, Z. Kockler, R. Evans, B.D. Downing, A. Malkova, Single-strand annealing between inverted DNA repeats: Pathway choice, participating proteins, and genome destabilizing consequences, PLoS Genet. 14 (2018) 1–29, doi:10.1371/journal.pgen.1007543.
[44] L.J. Marinelli, G.F. Hatfull, M. Piuri, Recombineering A powerful tool for modification of bacteriophage genomes, 2 (2012) 5–14, doi:10.4161/bact.18778.
[45] K.C. Murphy, λ Recombination and Recombineering, EcoSal Plus 7 (2016), doi:10.1128/ECOSALPLUS.ESP-0011-2015.
[46] S.K. Sharan, L.C. Thomason, S.G. Kuznetsov, D.L. Court, Recombineering: A homologous recombination-based method of genetic engineering, Nat. Protoc. 4 (2009) 206–223, doi:10.1038/nprot.2008.227.
[47] R.J. Thresher, A.M. Makhov, S.D. Hall, R. Kolodner, J.D. Griffith, Electron microscopic visualization of RecT protein and its complexes with DNA, J. Mol. Biol. 254 (1995) 364–371, doi:10.1006/JMBI.1995.0623.
[48] E. Kmiec, W.K. Holloman, Beta protein of bacteriophage lambda promotes renaturation of DNA, J. Biol. Chem. 256 (1981) 12636–12639, doi:10.1016/s0021-9258(18)42938-9.
[49] G. Karakousis, N. Ye, Z. Li, S.K. Chiu, G. Reddy, C.M. Radding, The beta protein of phage λ binds preferentially to an intermediate in DNA renaturation, J. Mol. Biol. 276 (1998) 721–731, doi:10.1006/jmbi.1997.1573.
[50] S. Weerasooriya, K.A. DiScipio, A.S. Darwish, P. Bai, S.K. Weller, Herpes simplex virus 1 ICP8 mutant lacking annealing activity is deficient for viral DNA replication, Proc. Natl. Acad. Sci. U. S. A. 116 (2019) 1033–1042, doi:10.1073/pnas.1817642116.
[51] M.E. O'Donnell, P. Elias, B.E. Funnell, I.R. Lehman, Interaction between the DNA polymerase and single-stranded DNA-binding protein (infected cell protein 8) of herpes simplex virus 1, J. Biol. Chem. 262 (1987) 4260–4266, doi:10.1016/S0021-9258(18)61341-9.
[52] G.M. Weinstock, K. McEntee, I.R. Lehman, ATP-dependent renaturation of DNA catalyzed by the recA protein of Escherichia coli, Proc. Natl. Acad. Sci. 76 (1979) 126–130, doi:10.1073/PNAS.76.1.126.
[53] C.M. Radding, J. Rosenzweig, F. Richards, E. Cassuto, Separation and characterization of exonuclease, β protein, and a complex of both, J. Biol. Chem. 246 (1971) 2510–2512, doi:10.1016/s0021-9258(18)62317-8.
[54] J.P.P. Muyrers, Y. Zhang, F. Buchholz, A.F. Stewart, RecE/RecT and Redα/Redβ initiate double-stranded break repair by specifically interacting with their respective partners, Genes Dev. 14 (2000) 1971–1982, doi:10.1101/gad.14.15.1971.
[55] A. Lo Piano, M.I. Martínez-Jiménez, L. Zecchi, S. Ayora, Recombination-dependent concatemeric viral DNA replication, Virus Res. 160 (2011) 1–14.
[56] A. Kuzminov, Recombinational Repair of DNA Damage in Escherichia coli and Bacteriophage λ, Microbiol. Mol. Biol. Rev. 63 (1999) 751–813 1999/ASSET/BC82C2F5-9E89-4251-A63F-668EF821FC51/ASSETS/GRAPHIC/MR0490033038.JPEG, doi:10.1128/MMBR.63.4.751-813.
[57] D. Roy, K.T. Huguet, F. Grenier, V. Burrus, IncC conjugative plasmids and SXT/R391 elements repair double-strand breaks caused by CRISPR-Cas during conjugation, Nucleic Acids Res. 48 (2020) 8815–8827.
[58] J.T. Martinsohn, M. Radman, M.A. Petit, The λ Red Proteins Promote Efficient Recombination between Diverged Sequences: Implications for Bacteriophage Genome Mosaicism, PLoS Genet. 4 (2008) e1000065, doi:10.1371/JOURNAL.PGEN.1000065.
[59] M. De Paepe, G. Hutinet, O. Son, J. Amarir-Bouhram, S. Schbath, M.A. Petit, Temperate Phages Acquire DNA from Defective Prophages by Relaxed Homologous Recombination: The Role of Rad52-Like Recombinases, PLoS Genet. 10 (2014) e1004181, doi:10.1371/JOURNAL.PGEN.1004181.
[60] M.M. Stahl, L. Thomason, A.R. Poteete, T. Tarkowski, A. Kuzminov, F.W. Stahl, Annealing vs. invasion in phage lambda recombination, Genetics 147 (1997) 961–977, doi:10.1093/GENETICS/147.3.961.
[61] J. Fishman-Lobell, N. Rudin, J.E. Haber, Two Alternative Pathways of Double-Strand Break Repair That Are Kinetically Separable and Independently Modulated, 1992.12 (2023) 1292–1303. https://doi.org/10.1128/MCB.12.3.1292-1303.1992.
[62] M. Maresca, A. Erler, J. Fu, A. Friedrich, Y. Zhang, A.F. Stewart, Single-stranded heteroduplex intermediates in λ Red homologous recombination, BMC Mol. Biol. 11 (2010), doi:10.1186/1471-2199-11-54.
[63] J.A. Mosberg, M.J. Lajoie, G.M. Church, Lambda Red Recombineering in Escherichia coli Occurs Through a Fully Single-Stranded Intermediate, Genetics 186 (2010) 791–799, doi:10.1534/GENETICS.110.120782.
[64] H.M. Ellis, D. Yu, T. DiTizio, D.L. Court, High efficiency mutagenesis, repair, and engineering of chromosomal DNA using single-stranded oligonucleotides, Proc. Natl. Acad. Sci. 98 (2001) 6742–6746, doi:10.1073/PNAS.121164898.
[65] A.R. Poteete, Involvement of DNA replication in phage lambda Red-mediated homologous recombination, Mol. Microbiol. 68 (2008) 66–74, doi:10.1111/J.1365-2958.2008.06133.X.
[66] J. Dapprich, Single-Molecule DNA Digestion by Lambda-Exonuclease, Cytometry 36 (1999) 163–168, doi:10.1002/(SICI)1097-0320(19990701)36:3.
[67] A.M. Van Oijen, P.C. Blainey, D.J. Crampton, C.C. Richardson, T. Ellenberger, X.S. Xie, Single-molecule kinetics of λ exonuclease reveal base dependence and dynamic disorder, Science 301 (2003) (1979) 1235–1238, doi:10.1126/science.1084387.
[68] F.W. Stahl, Recombination in phage λ: one geneticist's historical perspective, Gene 223 (1998) 95–102, doi:10.1016/S0378-1119(98)00246-7.
[69] B.J. Caldwell, E. Zakharova, G.T. Filsinger, T.M. Wannier, J.P. Hempfling, L. Chun-Der, D. Pei, G.M. Church, C.E. Bell, Crystal structure of the Red C-terminal domain in complex with Exonuclease reveals an unexpected homology with Orf and an interaction with Escherichia coli single stranded DNA binding protein, Nucleic Acids Res. 47 (2019) 1950–1963, doi:10.1093/nar/gky1309.
[70] S.V. Rajagopala, S. Casjens, P. Uetz, The protein interaction map of bacteriophage lambda, BMC Microbiol. 11 (2011) 213, doi:10.1186/1471-2180-11-213.
[71] P.E. Boehmer, I.R. Lehman, Physical interaction between the herpes simplex virus 1 origin-binding protein and single-stranded DNA-binding protein ICP8, Proc. Natl. Acad. Sci. 90 (1993) 8444–8448, doi:10.1073/pnas.90.18.8444.
[72] M. Olesky, E.E. McNamee, C. Zhou, T.J. Taylor, D.M. Knipe, Evidence for a direct interaction between HSV-1 ICP27 and ICP8 proteins, Virology 331 (2005) 94–105, doi:10.1016/j.virol.2004.10.003.
[73] A. Weissbach, D. Korn, The Effect of Lysogenic Induction on the Deoxyribonucleases of Escherichia coli K12λ, J. Biol. Chem. 237 (1962) PC3312–PC3314, doi:10.1016/S0021-9258(18)50170-8.
[74] C.M. Radding, D.C. Shreffler, Regulation of λ exonuclease: II. Joint regulation of exonuclease and a new λ antigen, J. Mol. Biol. 18 (1966) 251–261, doi:10.1016/S0022-2836(66)80244-9.
[75] J.W. Little, An Exonuclease Induced by Bacteriophage λ II. Nature of the Enzymatic Reaction, J. Biol. Chem. 242 (1967) 679–686, doi:10.1016/S0021-9258(18)96258-7.
[76] A.J. Clark, Toward a Metabolic Interpretation of Genetic Recombination of E. coli and its Phages, 25 (1971) 437–464, doi:10.1146/annurev.mi.25.100171.002253.
[77] S.D. Hall, M.F. Kane, R.D. Kolodner, Identification and characterization of the Escherichia coli RecT protein, a protein encoded by the RecE region that promotes renaturation of homologous single-stranded DNA, J. Bacteriol. 175 (1993) 277–287, doi:10.1128/jb.175.1.277-287.1993.

[78] R. Kovall, B.W. Matthews, Toroidal structure of λ-exonuclease, Science 277 (1997) (1979) 1824–1827, doi:10.1126/science.277.5333.1824.

[79] S.I. Passy, X. Yu, Z. Li, C.M. Radding, E.H. Egelman, Rings and filaments of β protein from bacteriophage λ suggest a superfamily of recombination proteins, Proc. Natl. Acad. Sci. U.S.A. 96 (1999) 4279–4284, doi:10.1073/pnas.96.8.4279.

[80] J. Zhang, X. Xing, A.B. Herr, C.E. Bell, Crystal Structure of E. coli RecE Protein Reveals a Toroidal Tetramer for Processing Double-Stranded DNA Breaks, Structure 17 (2009) 690–702, doi:10.1016/J.STR.2009.03.008.

[81] T.P. Newing, J.L. Brewster, L.J. Fitschen, J.C. Bouwer, N.P. Johnston, H. Yu, G. Tolun, Redβ177 annealase structure reveals details of oligomerization and λ Red-mediated homologous DNA recombination, Nat. Commun. 13 (2022) 5649, doi:10.1038/s41467-022-33090-6.

[82] J. Zhang, K.A. McCabe, C.E. Bell, Crystal structures of λ exonuclease in complex with DNA suggest an electrostatic ratchet mechanism for processivity, Proc. Natl. Acad. Sci. U.S.A. 108 (2011) 11872–11877, doi:10.1073/pnas.1103467108.

[83] J. Zhang, X. Pan, C.E. Bell, Crystal structure of λ exonuclease in complex with DNA and Ca2+, Biochemistry 53 (2014) 7415–7425, doi:10.1021/bi501155q.

[84] B.J. Caldwell, A.S. Norris, C.F. Karbowski, A.M. Wiegand, V.H. Wysocki, C.E. Bell, Structure of a RecT/Redβ family recombinase in complex with a duplex intermediate of DNA annealing, Nat. Commun. 13 (2022) 1–14, doi:10.1038/s41467-022-35572-z.

[85] A. Erler, S. Wegmann, C. Elie-Caille, C.R. Bradshaw, M. Maresca, R. Seidel, B. Habermann, D.J. Muller, A.F. Stewart, Conformational Adaptability of Redβ during DNA Annealing and Implications for Its Structural Relationship with Rad52, J. Mol. Biol. 391 (2009) 586–598, doi:10.1016/j.jmb.2009.06.030.

[86] R.G. Higuchi, H. Ochman1, Production of single-stranded DNA templates by exonuclease digestion following the polymerase Production of single-stranded DNA templates by exonuclease digestion following the polymerase chain reaction, Nucleic Acids Res. 17 (1989) 5865, doi:10.1093/nar/17.14.5865.

[87] H.S. Rhee, B.F. Pugh, Comprehensive Genome-wide Protein-DNA Interactions Detected at Single-Nucleotide Resolution, Cell 147 (2011) 1408–1419, doi:10.1016/j.cell.2011.11.013.

[88] Z.J. Liu, L.Y. Yang, T.C. Lu, Y.Q. Liang, M.M. Liu, G.X. Zhong, X.H. Lin, P.F. Huang, J.Y. Chen, A zero-background electrochemical DNA sensor coupling ligase chain reaction with lambda exonuclease digestion for CYP2C19*2 allele genotyping in clinical samples, Sens. Actuators B Chem. 368 (2022) 132096, doi:10.1016/j.snb.2022.132096.

[89] L.L. Xu, W. Zhao, J. Pu, S. Wang, S. Liu, H. Li, R. Yu, A Pax-5a gene analysis approach enabled by selective digestion with lambda exonuclease, Anal. Methods 14 (2022) 2415–2422, doi:10.1039/D2AY00469K.

[90] L. Cui, Y. Li, M. Lu, B. Tang, C. yang Zhang, An ultrasensitive electrochemical biosensor for polynucleotide kinase assay based on gold nanoparticle-mediated lambda exonuclease cleavage-induced signal amplification, Biosens. Bioelectron. 99 (2018) 1–7, doi:10.1016/j.bios.2017.07.028.

[91] J. Sun, C. Li, Y. Hu, Y. Ding, T. Wu, A structure change-induced fluorescent biosensor for uracil-DNA glycosylase activity detection based on the substrate preference of Lambda exonuclease, Talanta 243 (2022) 123350, doi:10.1016/j.talanta.2022.123350.

[92] A. Lopes, J. Amarir-Bouhram, G. Faure, M.A. Petit, R. Guerois, Detection of novel recombinases in bacteriophage genomes unveils Rad52, Rad51 and Gp2.5 remote homologs, Nucleic Acids Res. 38 (2010) 3952–3962, doi:10.1093/nar/gkq096.

[93] W. Yang, W.Y. Chen, H. Wang, J.W.S. Ho, J.D. Huang, P.C.Y. Woo, S.K.P. Lau, K.Y. Yuen, Q. Zhang, W. Zhou, M. Bartlam, R.M. Watt, Z. Rao, Structural and functional insight into the mechanism of an alkaline exonuclease from Laribacter hongkongensis, Nucleic Acids Res. 39 (2011) 9803–9819, doi:10.1093/nar/gkr660.

[94] I.V. Shevelev, U. Hübscher, The 3′–5′ exonucleases, Nat. Rev. Mol. Cell Biol. 3 (2002) 364–376 200253, doi:10.1038/nrm804.

[95] S.H. Mueller, L.J. Fitschen, A. Shirbini, S.M. Hamdan, L.M. Spenkelink, A.M. van Oijen, Rapid single-molecule characterisation of enzymes involved in nucleic-acid metabolism, Nucleic Acids Res. 1 (2022) e5, doi:10.1093/nar/gkac949.

[96] X. Pan, C.E. Smith, J. Zhang, K.A. McCabe, J. Fu, C.E. Bell, A Structure-Activity Analysis for Probing the Mechanism of Processive Double-Stranded DNA Digestion by λ Exonuclease Trimers, Biochemistry 54 (2015) 6139–6148, doi:10.1021/acs.biochem.5b00707.

[97] X. Pan, J. Yan, A. Patel, V.H. Wysocki, C.E. Bell, Mutant poisoning demonstrates a nonsequential mechanism for digestion of double-stranded DNA by λ exonuclease trimers, Biochemistry 54 (2015) 942–951, doi:10.1021/bi501431w.

[98] G. Tolun, R.S. Myers, A real-time DNase assay (ReDA) based on PicoGreen ® fluorescence, Nucleic Acids Res. 31 (2003) e111, doi:10.1093/nar/gng111.

[99] P.G. Mitsis, J.G. Kwagh, Characterization of the interaction of lambda exonuclease with the ends of DNA, Nucleic Acids Res. 27 (1999) 3057–3063, doi:10.1093/NAR/27.15.3057.

[100] K. Subramanian, W. Rutvisuttinunt, W. Scott, R.S. Myers, The enzymatic basis of processivity in λ exonuclease, Nucleic Acids Res. 31 (2003) 1585–1596, doi:10.1093/NAR/GKG266.

[101] J. van Oostrum, J.L. White, R.M. Burnett, Isolation and crystallization of λ exonuclease, Arch. Biochem. Biophys. 243 (1985) 332–337, doi:10.1016/0003-9861(85)90510-7.

[102] W. Yang, J.Y. Lee, M. Nowotny, Making and Breaking Nucleic Acids: Two-Mg2+-Ion Catalysis and Substrate Specificity, Mol. Cell 22 (2006) 5–13, doi:10.1016/J.MOLCEL.2006.03.013.

[103] A. Pingoud, M. Fuxreiter, V. Pingoud, W. Wende, Type II restriction endonucleases: Structure and mechanism, Cell. Mol. Life Sci. 62 (2005) 685–707, doi:10.1007/S00018-004-4513-1/METRICS.

[104] K. Steczkiewicz, A. Muszewska, L. Knizewski, L. Rychlewski, K. Ginalski, Sequence, structure and functional diversity of PD-(D/E)XK phosphodiesterase superfamily, Nucleic Acids Res. 40 (2012) 7016–7045, doi:10.1093/NAR/GKS382.

[105] X. Pan, C.E. Smith, J. Zhang, K.A. McCabe, J. Fu, C.E. Bell, A Structure-Activity Analysis for Probing the Mechanism of Processive Double-Stranded DNA Digestion by λ Exonuclease Trimers, Biochemistry 54 (2015) 6139–6148, doi:10.1021/ACS.BIOCHEM.5B00707.

[106] G. Tolun, More than the sum of its parts: physical and mechanistic coupling in the phage lambda red recombinase, (2007).

[107] S.D. Hall, R.D. Kolodner, Homologous pairing and strand exchange promoted by the Escherichia coli RecT protein, Proc. Natl. Acad. Sci. U.S.A. 91 (1994) 3205–3209, doi:10.1073/pnas.91.8.3205.

[108] J. Fu, X. Bian, S. Hu, H. Wang, F. Huang, P.M. Seibert, A. Plaza, L. Xia, R. Müller, A.F. Stewart, Y. Zhang, Full-length RecE enhances linear-linear homologous recombination and facilitates direct cloning for bioprospecting, Nat. Biotechnol. 30 (2012) 440–446, doi:10.1038/nbt.2183.

[109] T.S. Vellani, R.S. Myers, Bacteriophage SPP1 Chu is an alkaline exonuclease in the SynExo family of viral two-component recombinases, J. Bacteriol. 185 (2003) 2465–2474, doi:10.1128/JB.185.8.2465-2474.2003.

[110] I.N. Wang, Lysis Timing and Bacteriophage Fitness, Genetics 172 (2006) 17–26, doi:10.1534/genetics.105.045922.

[111] B. Gibson, D.J. Wilson, E. Feil, A. Eyre-Walker, The distribution of bacterial doubling times in the wild, Proc. R. Soc. B 285 (2018), doi:10.1098/rspb.2018.0789.

[112] K.C. Murphy, The λ Gam Protein Inhibits RecBCD Binding to dsDNA Ends, J. Mol. Biol. 371 (2007) 19–24, doi:10.1016/J.JMB.2007.05.085.

[113] A.E. Karu, Y. Sakaki, H. Echols, S. Linn, The gamma protein specified by bacteriophage lambda. Structure and inhibitory activity for the recBC enzyme of Escherichia coli, J. Biol. Chem. 250 (1975) 7377–7387, doi:10.1016/S0021-9258(19)40955-1.

[114] K.C. Murphy, Lambda Gam protein inhibits the helicase and chi-stimulated recombination activities of Escherichia coli RecBCD enzyme, J. Bacteriol. 173 (1991) 5808–5821, doi:10.1128/JB.173.18.5808-5821.1991.

[115] L.W. Enquist, A.M. Skalka, Replication of Bacteriophage λ DNA Dependent on the Function of Host and Viral Genes: I. Interaction of red, gam and rec, J. Mol. Biol. 75 (1972) 185–212, doi:10.1016/0022-2836(73)90016-8.

[116] J.L. Brewster, G. Tolun, Half a century of bacteriophage lambda recombinase: In vitro studies of lambda exonuclease and Red-beta annealase, IUBMB Life 72 (2020) 1622–1633, doi:10.1002/iub.2343.

[117] M. Wilkinson, L. Troman, W.A. Wan Nur Ismah, Y. Chaban, M.B. Avison, M.S. Dillingham, D.B. Wigley, Structural basis for the inhibition of RecBCD by Gam and its synergistic antibacterial effect with quinolones, eLife 5 (2016), doi:10.7554/ELIFE.22963.

[118] R. Court, N. Cook, K. Saikrishnan, D. Wigley, The Crystal Structure of λ-Gam Protein Suggests a Model for RecBCD Inhibition 271 (2007) 25–33, doi:10.1016/j.jmb.2007.05.037.

[119] B.J. Caldwell, A. Norris, E. Zakharova, C.E. Smith, C.T. Wheat, D. Choudhary, M. Sotomayor, V.H. Wysocki, C.E. Bell, Oligomeric complexes formed by Redβ single strand annealing protein in its different DNA bound states, Nucleic Acids Res. 49 (2021) 3441–3460, doi:10.1093/nar/gkab125.

[120] K. Steczkiewicz, E. Prestel, E. Bidnenko, A.K. Szczepankowska, Expanding Diversity of Firmicutes Single-Strand Annealing Proteins: A Putative Role of Bacteriophage-Host Arms Race, Front. Microbiol. 12 (2021) 644622, doi:10.3389/fmicb.2021.644622.

[121] A.J. Hernandez, C.C. Richardson, Gp2.5, the multifunctional bacteriophage T7 single-stranded DNA binding protein, Semin. Cell Dev. Biol. 86 (2019) 92–101, doi:10.1016/J.SEMCDB.2018.03.018.

[122] R.D. Shereda, A.G. Kozlov, T.M. Lohman, M.M. Cox, J.L. Keck, SSB as an Organizer/Mobilizer of Genome Maintenance Complexes, 43 (2008) 289–318, doi:10.1080/10409230802341296.

[123] A.S. Darwish, L.M. Grady, P. Bai, S.K. Weller, ICP8 Filament Formation Is Essential for Replication Compartment Formation during Herpes Simplex Virus Infection, J. Virol. 90 (2016) 2561–2570, doi:10.1128/JVI.02854-15.

[124] A.M. Makhov, A. Sen, X. Yu, M.N. Simon, J.D. Griffith, E.H. Egelman, The Bipolar Filaments Formed by Herpes Simplex Virus Type 1 SSB/Recombination Protein (ICP8) Suggest a Mechanism for DNA Annealing, J. Mol. Biol. 386 (2009) 273–279, doi:10.1016/J.JMB.2008.12.059.

[125] M. Ander, S. Subrmaniam, K. Fahmy, A.F. Stewart, E. Schäffer, A single-strand annealing protein clamps DNA to detect and secure homology, PLoS Biol. 13 (2015) e1002213.

[126] W. Kagawa, H. Kurumizaka, R. Ishitani, S. Fukai, O. Nureki, T. Shibata, S. Yokoyama, Crystal Structure of the Homologous-Pairing Domain from the Human Rad52 Recombinase in the Undecameric Form, Mol. Cell 10 (2002) 359–371, doi:10.1016/S1097-2765(02)00587-7.

[127] M.R. Singleton, L.M. Wentzell, Y. Liu, S.C. West, D.B. Wigley, Structure of the single-strand annealing domain of human RAD52 protein, Proc. Natl. Acad. Sci. U.S.A. 99 (2002) 13492–13497, doi:10.1073/pnas.212449899.

[128] M. Saotome, K. Saito, K. Onodera, H. Kurumizaka, W. Kagawa, Structure of the human DNA-repair protein RAD52 containing surface mutations, Acta Crystallogr. F Struct. Biol. Commun. 72 (2016) 598–603, doi:10.1107/S2053230x1601027X.

[129] C. Kinoshita, Y. Takizawa, M. Saotome, S. Ogino, H. Kurumizaka, W. Kagawa, C.W. Kagawa, H. Kurumizaka, The cryo-EM structure of full-length RAD52 protein contains an undecameric ring, FEBS Open Bio 13 (2023) 408–418, doi:10.1002/2211-5463.13565.

[130] A. Al-Fatlawi, M. Schroeder, A.F. Stewart, The Rad52 SSAP superfamily and new insights into homologous recombination, Commun. Biol. 6 (2023) 87, doi:10.1038/s42003-023-04476-z.

[131] A.Z. Stasiak, E. Larquet, A. Stasiak, S. Müller, A. Engel, E. Van Dyck, S.C. West, E.H. Egelman, The human Rad52 protein exists as a heptameric ring, Curr. Biol. 10 (2000) 337–340, doi:10.1016/S0960-9822(00)00385-7.
[132] M. Mapelli, S. Panjikar, P.A. Tucker, 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 (2005) 2990–2997, doi:10.1074/jbc.M406780200.
[133] M. Valledor, R.S. Myers, P.C. Schiller, Herpes ICP8 protein stimulates homologous recombination in human cells, PLoS One 13 (2018) e0200955, doi:10.1371/journal.pone.0200955.
[134] A.J. Schumacher, K.N. Mohni, Y. Kan, E.A. Hendrickson, J.M. Stark, S.K. Weller, The HSV-1 Exonuclease, UL12, Stimulates Recombination by a Single Strand Annealing Mechanism, PLoS Pathog. 8 (2012) e1002862, doi:10.1371/journal.ppat.1002862.
[135] M. Gao, D.M. Knipe, Potential role for herpes simplex virus ICP8 DNA replication protein in stimulation of late gene expression, J. Virol. 65 (1991) 2666–2675, doi:10.1128/JVI.65.5.2666-2675.1991.
[136] M. Gao, D.M. Knipe, Genetic evidence for multiple nuclear functions of the herpes simplex virus ICP8 DNA-binding protein, J. Virol. 63 (1989) 5258–5267, doi:10.1128/JVI.63.12.5258-5267.1989.
[137] T. Paysan-Lafosse, M. Blum, S. Chuguransky, T. Grego, B.L. Pinto, G.A. Salazar, M.L. Bileschi, P. Bork, A. Bridge, L. Colwell, J. Gough, D.H. Haft, I. Letunić, A. Marchler-Bauer, H. Mi, D.A. Natale, C.A. Orengo, A.P. Pandurangan, C. Rivoire, C.J.A. Sigrist, I. Sillitoe, N. Thanki, P.D. Thomas, S.C.E. Tosatto, C.H. Wu, A. Bateman, InterPro in 2022, Nucleic Acids Res. 51 (2023) D418–D427, doi:10.1093/nar/gkac993.
[138] K. Katoh, D.M. Standley, MAFFT Multiple Sequence Alignment Software Version 7: Improvements in Performance and Usability, Mol. Biol. Evol. 30 (2013) 772–780, doi:10.1093/molbev/mst010.
[139] L. Fu, B. Niu, Z. Zhu, S. Wu, W. Li, CD-HIT: accelerated for clustering the next-generation sequencing data, Bioinformatics 28 (2012) 3150–3152, doi:10.1093/bioinformatics/bts565.
[140] A.M. Waterhouse, J.B. Procter, D.M.A. Martin, M. Clamp, G.J. Barton, Jalview Version 2—A multiple sequence alignment editor and analysis workbench, Bioinformatics 25 (2009) 1189–1191, doi:10.1093/bioinformatics/btp033.
[141] E.F. Pettersen, T.D. Goddard, C.C. Huang, E.C. Meng, G.S. Couch, T.I. Croll, J.H. Morris, T.E. Ferrin, U.C.S.F. ChimeraX, Structure visualization for researchers, educators, and developers, Protein Sci. 30 (2021) 70–82, doi:10.1002/PRO.3943.
[142] J. Pei, N.V. Grishin, AL2CO: calculation of positional conservation in a protein sequence alignment, Bioinformatics 17 (2001) 700–712, doi:10.1093/bioinformatics/17.8.700.
[143] M. Mirdita, K. Schütze, Y. Moriwaki, L. Heo, S. Ovchinnikov, M. Steinegger, ColabFold: making protein folding accessible to all, Nat. Methods 19 (2022) 679–682, doi:10.1038/s41592-022-01488-1.