We are IntechOpen, the world's leading publisher of Open Access books Built by scientists, for scientists

6,900
Open access books available

185,000
International authors and editors

200M
Downloads

Our authors are among the

154
Countries delivered to

TOP 1%
most cited scientists

12.2%
Contributors from top 500 universities



WEB OF SCIENCE™

Selection of our books indexed in the Book Citation Index in Web of Science™ Core Collection (BKCI)

Interested in publishing with us?
Contact book.department@intechopen.com

Numbers displayed above are based on latest data collected.
For more information visit www.intechopen.com



Chapter

The Unusual Linear Plasmid Generating Systems of Prokaryotes

Sophie E. Knott, Sarah A. Milsom and Paul J. Rothwell

Abstract

Linear DNA is vulnerable to exonuclease degradation and suffers from genetic loss due to the end replication problem. Eukaryotes overcome these problems by locating repetitive telomere sequences at the end of each chromosome. In humans and other vertebrates this noncoding terminal sequence is repeated between hundreds and thousands of times, ensuring important genetic information is protected. In most prokaryotes, the end-replication problem is solved by utilizing circular DNA molecules as chromosomes. However, some phage and bacteria do store genetic information in linear constructs, and the ends of these structures form either invertrons or hairpin telomeres. Hairpin telomere formation is catalyzed by a protelomerase, a unique protein that modifies DNA by a two-step transesterification reaction, proceeding via a covalent protein bound intermediate. The specifics of this mechanism are largely unknown and conflicting data suggests variations occur between different systems. These proteins, and the DNA constructs they produce, have valuable applications in the biotechnology industry. They are also an essential component of some human pathogens, an increased understanding of how they operate is therefore of fundamental importance. Although this review will focus on phage encoded protelomerase, protelomerases found from *Agrobacterium* and *Borellia* will be discussed in terms of mechanism of action.

Keywords: protelomerase, telomere resolvase, linear plasmid, doggybone DNA, Touchlight Genetics

1. Introduction

The study of DNA, its structure and how it is replicated has been intensifying since the 1900s. Recent advances in DNA sequencing, bioinformatics and high-resolution imaging has increased our understanding of the variations that exist between different DNA replication systems. In general, the genetic material of bacterial cells is in the form of circular DNA molecules. Infecting bacteriophage may integrate their DNA into the host genome, or maintain it independently as a viral episome; usually this plasmid is also circular. These structures have no free ends and are therefore not susceptible to exonuclease degradation and do not suffer from the end replication problem, whereby genetic material at the tip of a chromosome is lost during each round of replication. However, some prokaryotic cells have been identified as harboring closed linear DNA chromosomes. The ends of these structures

are protected by either invertron or hairpin telomeres. Invertron telomeres consist of inverted terminal repeats and covalently attached capping proteins, essential for priming DNA replication. These structures are distinct from hairpin telomeres, which have covalently closed hairpin ends.

The first linear genome of prokaryotes was obtained in 1964, when V. Ravin isolated the *Escherichia coli* phage N15 [1]. The genetic material of this phage is unusual, because it is not maintained as an independent circular entity nor is it integrated into the host genomic material. Instead, upon entry into the host, the N15 genome circularizes and is then processed into linear structures by an atypical cutting and re-joining enzyme, called a protelomerase, or telomere resolvase. Since this discovery, protelomerases have also been characterised in *Klebsiella oxytoca* phage φKO2, *Yersinia enterocolitica* phage PY54, *Vibrio parahaemolyticus* phage VP882 [2] and *Halomonas aquamarina* phage ΦHAP-1.

In addition to phage, these unusual enzymes have been isolated from certain bacteria. The best studied being ResT from *Borrelia burgdorferi* [3], the causative agent of Lyme disease. Linear chromosomes are now described as a hallmark of *Borrelia* [4] and protelomerases have been purified from *B. hermsii*, *B. parkeri*, *B. recurrentis*, *B. turicatae*, and *B. anserine* [5]. More recently, they have also been discovered in cyanobacteria [6] and the plant pathogen *Agrobacterium tumefaciens C58*, which contains a circular and linear chromosome as well as circular plasmids [7]. These proteins are clearly more widespread than initially believed and it is likely future research into other prokaryotes will identify additional members of the family.

Although currently under debate, it has been suggested protelomerases are tyrosine-recombinase-like enzymes. It remains to be determined whether the bacterial protelomerases have their origin in phage as both, to some extent, share a common substrate recognition and a DNA cleavage/rejoining mechanism. In addition, some protelomerases have been identified as having roles in vivo unrelated to telomere resolution such as single strand annealing [8] and ATP-dependent helicase activities [9]. Further research into this important class of enzymes should help elucidate the significance of single strand annealing and DNA unwinding activities during closed linear chromosome replication.

Not only are protelomerases essential for the organisms in which they reside, but their unique functionality makes them valuable to the biotechnology industry. DNA constructs produced using protelomerase are being marketed by Lucigen as improved cloning vectors for highly repetitive sequences [10, 11]. Linear structures are not susceptible to supercoiling, thus making them more stable and less susceptible to genetic loss during replication [10]. Protelomerases can also be expressed in engineered *E. coli* cells to produce linear eukaryotic vectors that contain no bacterial sequences [12]. In addition, protelomerases are a central component to Touchlight Genetics' DNA amplification platform that produces large quantities of high-quality DNA using a cell-free process, for therapeutic and industrial applications [13, 14]. Unlike plasmid DNA, the incumbent technology for therapeutic applications, Touchlight's doggybone DNA (dbDNA) platform contains no extraneous bacterial DNA sequences. The resulting minimal vector has an improved safety profile from a regulatory perspective due to elimination of antibiotic resistance genes. The small amounts of plasmid DNA required for this in vitro manufacturing process makes the dbDNA process well suited to scale production of "difficult" structured or repetitive DNA sequences or constructs that cause cell toxicity. Linear, minimal vectors may be valuable as nonviral gene therapy vectors and as DNA vaccines, both modalities gaining increasing focus and investment in the biotechnology market.

There is high interest in the study of this protein family due to their utility and potential value for biotechnology. To date, research has largely focused on characterizing protelomerase recognition sequences, solving 3-dimensional structures and exploring the effects of protein mutations on activity. An improved understanding of how protelomerases function will enhance their value for applications in synthetic biology, and may provide the opportunity to invent new and novel applications.

2. Mechanism of telomere resolution

Despite the diversity of organisms in which protelomerases reside, important features have been identified that unify and define this class of protein. The protelomerase target site, denoted as *telRL*, is a palindromic sequence of double stranded DNA. The substrate differs between protelomerases and to date only ResT, the bacterial protelomerase from *Borrelia*, has been shown to have specificity for more than one target sequence [15]. All protelomerases are thought to function as a dimer and it is widely believed that none require the addition of cofactors such as ATP or divalent cations. However, it has been shown that concentrations of EDTA >10 mm inhibit the N15 protelomerase, TelN, and the sequence of this protein predicts a binding motif for divalent cations [16].

Current models propose that protelomerases bind nonspecifically to DNA and scan the sequence until finding the target site, or coming into contact with another monomer, at which point the protein immobilizes [17]. Immobilization occurs upon dimerization, whether this forms at the substrate target sequence or not. However, only when at the correct site will the reaction of telomere formation be catalyzed. This phenomenon can be observed in vitro, where a high concentration of TelK (over 400 nm) results in the condensation of DNA and inhibition of telomere formation [17]. Protelomerase concentration in vivo therefore, must be carefully controlled. This notion has been explored in phage N15, where negative control is used to regulate the levels of protein [18].

Protelomerases catalyzes a two-step transesterification reaction, and all are thought to initiate DNA cleavage using an active site tyrosine residue. This residue performs nucleophilic attack on the phosphodiester bond to form a 3′ covalently attached protein-DNA intermediate and a free 5′-OH. The protein bound intermediate is vital for avoiding deleterious double strand breaks and prevents the premature abortion of reactions [19]. The DNA cleavage reaction happens in a staggered formation 3-bp either side of the symmetrical target site center. This leaves a 6-nucleotide overhang that loops back and is ligated to form the covalently closed hairpin end. The DNA cleaving and re-joining reactions are isoenergetic and, in principle, each step in the reaction is reversible [19]. As DNA hairpins are unable to form complete base pairings [20], they are less stable than the starting material. In this case directionality is determined by the loop processing step. This part of the reaction is poorly understood and data available indicates conflicting mechanisms in different systems.

Figure 1a is a model for telomere resolution by the protelomerase TelK from phage φKO2. An interlocked protein dimer forms at the *telRL* site and induces a sharp, roughly 73°, bend in the DNA, which displaces its helical structure and buckles the base pairs between the scissile phosphates [21]. This is described as "spring loading"; the energy stored in the distorted DNA drives the reaction forward, enabling spontaneous hairpin formation and protein dimer separation [19]. The mechanisms proposed for the bacterial protelomerases, TelA and ResT

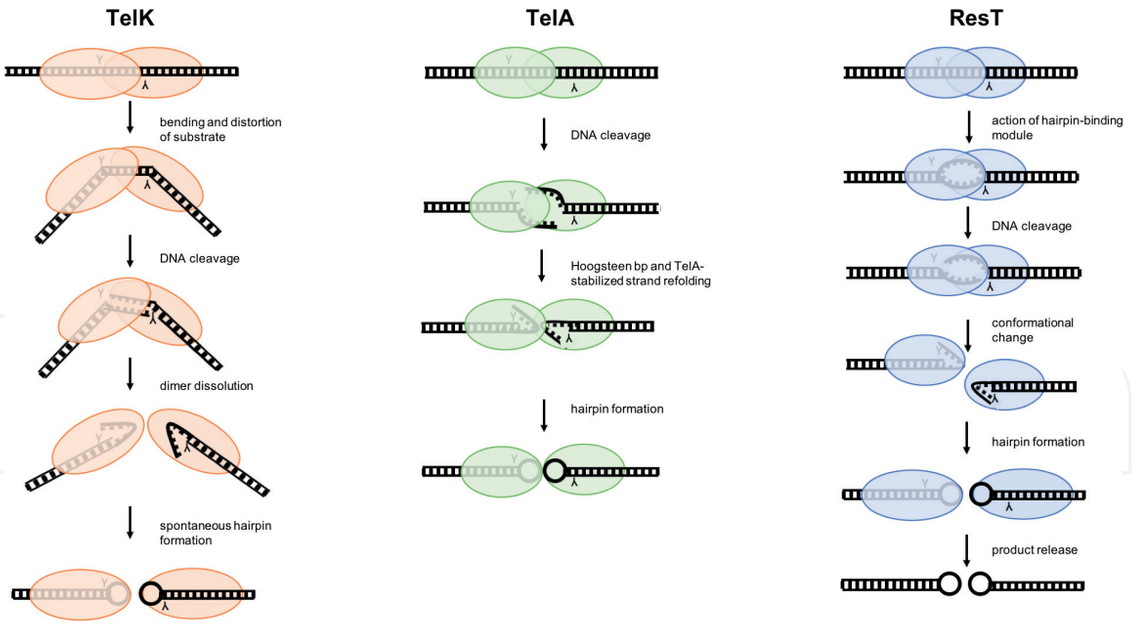


Figure 1.
Models of telomere resolution by TelK, TelA and ResT. (a) TelK monomers composed of N-terminal muzzle and C-terminal stirrup domains dimerise at the target site. This induces bending of the DNA, the spontaneous release of stored energy drives hairpin formation and dimer dissolution. (b) TelA cleaves the DNA and transient electrostatic interactions stabilize the transition state. Hairpin formation occurs within the protein dimer (c) ResT catalyzes telomere resolution with the aid of its hairpin-binding module. The final step of this reaction is product release, which is not observed for TelK or TelA [19].

(from *Agrobacterium tumefaciens C58* and *Borrelia*, respectively) are fundamentally different to that of TelK. In TelA and ResT reactions, strand refolding is enzyme-mediated, as opposed to spontaneous. A key element of the TelA mechanism is the refolding intermediate that exists before hairpin formation. This conformation is stabilized by multiple protein-DNA and DNA-DNA interactions, which drive the reaction forward by virtue of changes in binding energies. TelA binds even more strongly to the final hairpin product, thus favoring its formation. The mechanisms for TelK and TelA have been deduced from structures solved by X-ray crystallography [21, 22]. There is no structure of ResT and the mechanism proposed in **Figure 1c** is a result of research involving structure prediction, substrate modifications and protein mutations. In ResT catalyzed telomere resolution, the protein binds and distorts the DNA by underwinding at the dimer interface [19]. This is consistent with the observation that ResT has a hairpin-binding module, that presumably stabilizes the conformation of pre-hairpin DNA [23]. Hydrolysis of base pairs between the scissile phosphates promotes strand ejection following DNA cleavage. The exact mechanism of strand refolding is yet to be determined, but it is suggested to occur before dissolution of the dimer [24]. This concept of a "spring-loaded" pre-cleavage intermediate is analogous to that of TelK.

3. Substrate sequences

Identifying the natural target site of protelomerases is not straightforward. A logical strategy is to determine the nucleotide sequence of the resultant telomere and deduce from this the starting material. However, sequencing telomeres is notoriously difficult as the hairpin ends are incapable of ligating to the vector during sequencing library construction [7]. An adapted method has been used,

whereby a nuclease opens the closed ends to make them compliant for ligation [25]. This does not always give absolute results, but can provide predictions for the target sequences, which may be confirmed by in vitro studies [25].

In general, protelomerases are highly specific and only process one target sequence. The exception to this is ResT, which is far less stringent and can resolve nine different telomere sequences found in the *B. burgdorferi* group of bacteria [15]. A conserved feature among all protelomerases is the palindromic nature of their substrate, with one protein molecule binding either side of the axis of symmetry to form a dimer. Interestingly, the TATAAT sequence of telomeres from N15 and φKO2 is also found in *Borrelia*. The significance of this is unconfirmed, although it has been suggested the nucleotides are important for protelomerase recognition [15]. For ResT, substitution of this sequence abolishes telomere resolution and mutating it to TTTAAT reduces the initial rate significantly [15]. Mutating the 6th and 7th nucleotide of this sequence within the TelN recognition site also produces a substrate the protein cannot process [26]. Despite functioning in different systems, TelN and TelK process highly similar target sequences, both of which are shown in **Figure 2**. These sites differ in length, but are identical in the center, and both protelomerases are capable of resolving each other's natural substrate [27]. Given the high sequence similarity (86.9%) between TelN and TelK, this observation is not hugely surprising.

Comparison of the TelN and TelK *telRL* sequences, to the 42 base pair (bp) recognition site of the PY54 protelomerase indicates limited homology and this DNA cannot be processed by any of the other protelomerases [27]. However, Huang and colleagues found altering positions 15 and 16 of the PY54 target in the top strand, plus residues 28 and 27 of the bottom strand results in a substrate that is processed, although with limited efficiency, by TelK. They went on to suggest that TelN and TelK not only recognize these specific nucleotides, but also a cruciform DNA structure that is formed [27]. Although crystal structures of TelK have since discredited the suggestion that a cruciform structure is formed [21], this work is important in that it identifies the key nucleotides that are essential for telomere resolution by these enzymes.

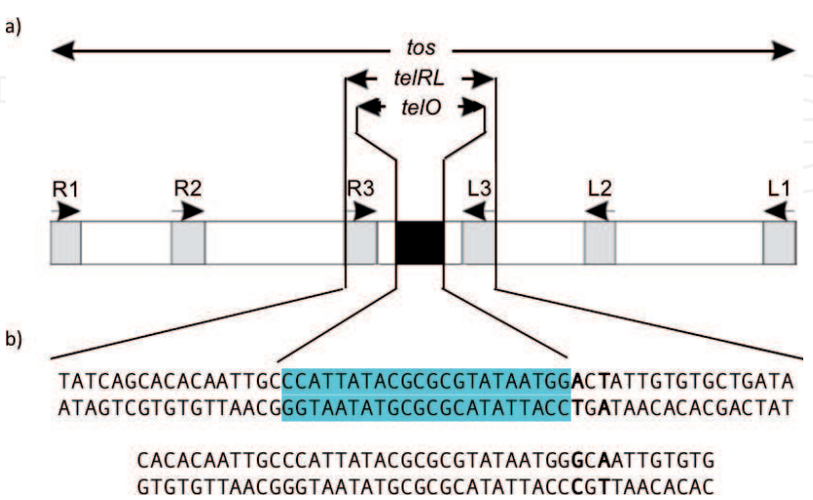


Figure 2.
Protelomerase recognition sequences. (a) The tos *site for TelN. Gray boxes indicate three regions of repeated sequences flanking the telRL site, which contains the central 22-bp telO site highlighted in cyan. Figure adapted from [16]. (b) The cognate sequence of TelN (top) and TelK (bottom). Both protelomerases are capable of processing each other's substrate. Two single point variations are shown in bold. In addition, the TelN natural substrate has six residues on each end.*

3.1 Minimal substrate

In vitro studies have also involved truncating target sites in order to identify the minimum sequence required for protelomerase binding and telomere resolution. To date, the minimal site identified that can be resolved, is a 26-bp substrate of TelA [7]. This was found by systematically deleting residues from both sides of the target sequence, until no product was produced. Similar studies have been performed on the TelN substrate. **Figure 2** shows the complete telomere occupancy (*tos*) site, which consists of a 56-bp palindromic sequence flanked by a series of inverted repeats. Initially, it was believed that *telO* is insufficient for processing by TelN, and the reaction requires the whole *telRL* site [16]. However, it has since been found that at greater TelN concentrations, roughly 50-fold higher than those required for *telRL*, the *telO* substrate is processed [26]. This indicates *telO* contains all the necessary elements for telomere resolution, but the protein requires additional sequence for binding and recognition. The binding affinity of TelN is greater still when the whole *tos* site is included in the substrate [16]. Experiments performed using ResT have explored whether the protelomerase is able to mediate cleavage on half a target site. When this half site was in a plasmid, the assay failed to produce reaction products, therefore suggesting dimer formation is essential for activity and the whole palindromic sequence is required [28].

4. Linear genomes

In order to further appreciate how protelomerases function, it is necessary to understand their role in relation to the whole phage or bacterial cell life-cycle. Bacteriophage N15 has been extensively characterised and **Figure 3** illustrates the different structures its genetic material forms upon infection of *E. coli*. The phage

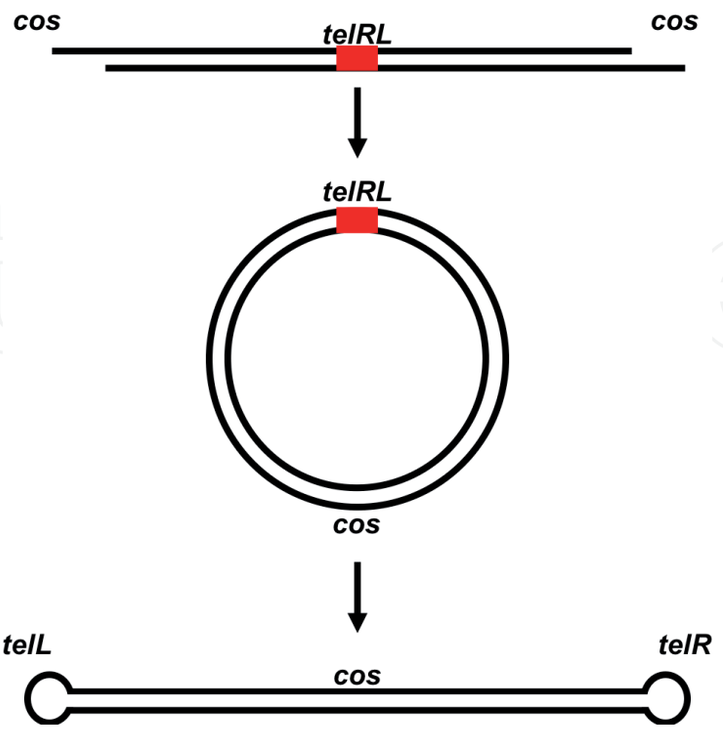


Figure 3.
Forms of bacteriophage N15 DNA. Having infected an E. coli *cell, the virion DNA circularizes, via complementary* cos *sites. Lytic or lysogenic replication can be initiated from the circular intermediate. Shown is the pathway for lysogeny, whereby the telRL is processed by protelomerases to form linear prophage DNA with covalently closed hairpin ends. Figure adapted from [1].*

DNA is a 46.4 kb chromosome that has two cohesive end sites (*cos*) consisting of 12-nucleotide overhangs at each 5′-end. These sites are complementary and can be ligated to form circular DNA. This circular intermediate then acts as the starting material for either lytic (not shown in **Figure 3**) or lysogenic development. During lysogenic development the *telRL* site is recognized and processed by protelomerases, this reaction forms a linear DNA structure with covalently closed ends. A similar genome arrangement has been identified for φKO2 [27], VP58.5 [29], VP882 [2] and PY54 [30]. These prophages all have cohesive (*cos*) ends that presumably enable the formation of similar structures as those described for N15. Interestingly, no *cos* site has been identified in the ΦHAP1 genome [31], therefore indicating a different mechanism of DNA packing. The discovery of terminase genes [31], suggests that headful packing may occur, whereby concatermeric DNA is packed into the phage capsid until it is full [32].

5. Bacteriophage N15 replication

Various models have been proposed to describe the replication and processing of linear DNA with hairpin telomeres [33]. Uncertainties arise about the specific mode of replication and whether it occurs uni- or bi-directionally. Other important factors that need to be determined are, where in the plasmid replication is initiated from and what the replication intermediates are. Bacteriophage N15 can be used as the model system to explore these questions, the general organization of its genome is shown in **Figure 4**. Genes have been largely identified by homology inferred from sequence similarity to other bacteriophages; mainly lambda, HK97 and HK002 [1]. The division between the left- and right-hand side of the N15 genome is marked by *telRL*. The left arm encodes structural proteins required for N15 head and tail assembly. The right-arm contains more unusual genes and only 10 of the 35 have identified homologs in other lambdoid phage [1]. These are therefore much harder to characterize, and it is yet to be determined how they all function during N15 replication.

RepA is the only gene essential for replication of prophage N15 DNA [18]. It encodes a large, multifunctional protein that has both primase and helicase activities [35]. Sequence alignments have highlighted regions of RepA with similarities to both plasmid and viral DNA replication proteins [1]. Most notably, the phage P4 alpha protein [1] also has combined primase and helicase activities [36]. Phage P4 replication occurs by a theta-mechanism [37]. The similarities between alpha protein and RepA, combined with studies measuring amplification rates of DNA markers [18], strongly suggests that typical bidirectional theta-replication also occurs in N15 prophage. The origin of replication (*ori*) resides within the *repA* gene [18], which is located closer to the left hairpin end of the plasmid.

The gene *telN* encodes a 71 kDa protein that has partial homology to integrases and an amino acid sequence characteristic of those that bind DNA as homo- or hetero- dimers [38, 39]. This was correctly identified as the protelomerase encoding

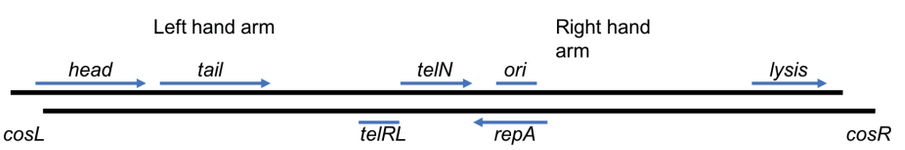


Figure 4.
The chromosome of bacteriophage N15. 46.4 kb double stranded DNA with 12-bp single stranded cohesive termini (cosL and cosR). Arrows indicate the direction of transcription and the telRL site divides the sequence into left- and right-hand arms. Adapted from [34].

gene and in 2000, Deneke and colleagues purified its protein product [16]. TelN is capable of processing the 56-bp *telRL* site in both linear and circular supercoiled DNA [16]. To decipher the mechanism of N15 genomic replication, mutants deficient in this protein have been created [40]. In protelomerase-deficient cells, unprocessed replicative intermediates accumulate, the structures of which have been characterised as circular head-to-head dimer molecules [40].

Figure 5 describes how these linear N15 constructs may be replicated and processed; it is consistent with the data cited above and proposes structures that have been validated by electron microscopy [18]. In pathway A, following replication of the *telL* site, TelN processes the DNA to create a Y-shaped structure. After duplication of *telR*, the right telomere is also modified to form the final linear product. Alternatively, in pathway B the whole DNA molecule has been replicated, producing a head-to-head circular dimer that is then resolved. Interestingly, this mechanism of replication is distinct from that described for eukaryotic replicons, even those with similar telomeric ends, therefore suggesting an independent evolution [18].

5.1 Lytic replication

A model of how N15 lytic replication could occur is proposed in pathway C of **Figure 2**. The DNA is duplicated and resolved into two circular monomers, as opposed to linear structures. These circular molecules are the starting material for subsequent cycles of amplification. This style of lytic replication is similar to that of phage lambda. This bacteriophage also circularizes its DNA upon entering the host cell, it has *cos* sites analogous to those of N15 [1]. Further similarities between N15 and lambda include: genome length, burst size, latent period, lysogenization frequency and phage particle and plaque morphology [38]. Their structural and packing proteins are also highly analogous, making it likely that N15 DNA packing follows a pathway similar to that of lambda [1]. It has even

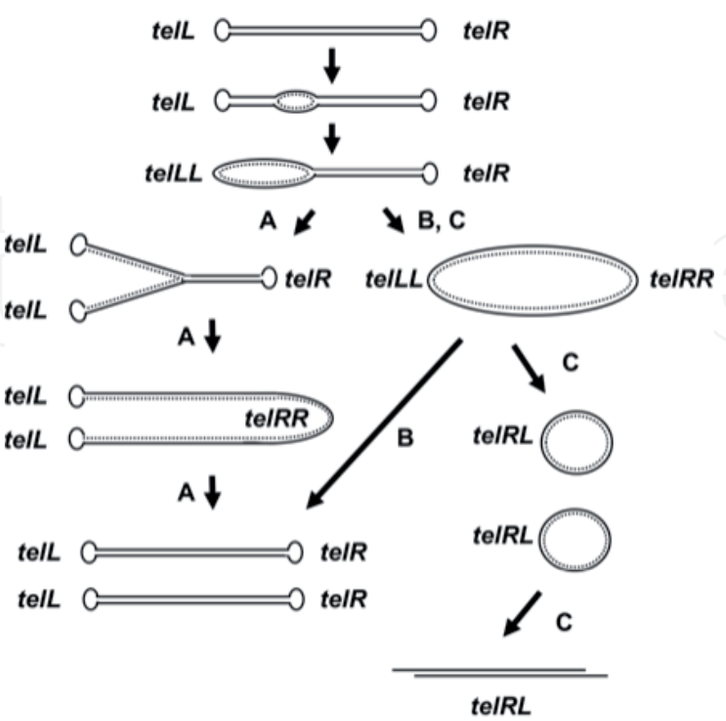


Figure 5.
Model for lytic and lysogenic replication of N15 linear prophage. Bi-directional theta-replication begins at the internal ori *site. A: Duplicated* telL *sites are processed before complete replication of genome. B: The template is completely duplicated prior to processing by TelN. C: Lytic replication, whereby circular monomers are produced that then undergo subsequent rounds of replication [18].*

been demonstrated that the N15 specific terminase can package lambda DNA with reasonable efficiency [41].

The key difference between N15 and lambda bacteriophage is that lambda integrates its DNA into the host genome, whereas N15 does not. Although protelomerases share some sequence homology with lambda integrases, and both appear to have comparable roles in helping establish prophage DNA, these proteins are not functional analogues. During lytic replication the lambda integrase is dispensable, in comparison the protelomerase of N15 is essential. This phenomenon has been proven by experiments showing N15 deficient in protelomerases are incapable of infecting *E. coli* cells [35], although why this is the case remains unclear. As the establishment of lytic growth requires the conversion of linear plasmid molecules to circular ones, it could be presumed a protelomerase mediated "telomere fusion" reaction occurs. However, TelK, which is highly analogous to TelN, is incapable of catalyzing this in vitro [21], it is therefore highly unlikely wild-type TelN is functioning in this way. The possibility of an unknown factor modifying the protelomerase and/or its target site to prevent the usual processing reaction cannot be ruled out. In lambda, Xis, assisted by the host factor Fis, is necessary to induce excision during induction of a lysogeny [42]. Potentially, an analogue could be encoded by one of the N15 late genes [35], although experimental evidence to support this theory is yet to be provided. However, it has been demonstrated that mutating histidine 415 of TelN to an alanine results in accumulation of circular head-to-tail monomers [35]. This histidine is important for catalytic activity and is believed to coordinate the scissile phosphate [16]. Interestingly, its mutation does not have the same effect as mutating the catalytic tyrosine 424, which acts as the nucleophile in telomere resolution. When this residue is changed to an alanine, accumulation of circular dimers does occur, but in this case, they are "head-to-head" as opposed to "head-to-tail" [35]. The significance of this observation is currently unknown. Given that TelN cannot be recycled [27], it has also been proposed that the protein's depletion will result in fewer linear molecules being produced and a natural accumulation of head-to-head dimers, which can then be processed to circular structures [35].

5.2 N15 as a model for the replication of other linear plasmids

To what extent can the model of bacteriophage N15 be extended to describe the replication of other replicons with hairpin ends? Genomic sequence analysis of phage encoded protelomerases reveal little overall sequence similarity [43]. However, the organization of functional domains is analogous and they appear to have conserved regulatory regions [43]. This would suggest a shared mechanism of plasmid replication and lysogeny control [43]. Virions of φKO2, VP58.5, VP882 and PY54 have cohesive ends [27, 29, 2], which facilitate circularization and enable the formation of similar structures as those described for N15. The absence of *cos* sites in the ΦHAP1 genome has already been discussed and suggests a different mechanism of DNA packing [31]. Importantly, these phages all have homologs of the N15 protelomerase and replication protein RepA. The genes encoding these proteins are found between the lysogeny control region and structural gene cluster, as is the case is N15 [43]. Although yet to be confirmed by in vitro studies, given these similarities it is sensible to suggest that replication of these linear phage plasmids follows a model comparable to that proposed for N15.

Further comparisons can be made between the suggested phage model and that of bacterial linear chromosome replication. *B. burgdorferi* has a linear chromosome [44], and it is replicated in a bi-directional manner to produce circular, head-to-head intermediates [45], which are then processed by telomere resolution [46]. Here replication is also initiated at an internal *ori* site and the protelomerase, ResT,

is known to be essential [47]. These findings indicate a shared fundamental mechanism of genomic replication between N15 and *Borrelia*. Nonetheless, discrepancies between the different systems have been highlighted. For one, in these bacterial cells the protelomerase is encoded, not on the same DNA construct it processes, but on a different circular plasmid, cp26 [3]. In addition, the possibility of *Borrelia* accessory factors influencing telomere resolution was suggested, following the observation that differential processing occurs in vitro compared to in vivo [15]. Potentially this could be a result of in vitro conditions not completely reconstructing those occurring in vivo. However, if correct it would indicate important discrepancies between how bacterial and phage protelomerases are regulated.

6. Structural data

6.1 X-ray crystallography

The X-ray structures described for both TelK and TelA have greatly enhanced our understanding of the protelomerase mechanism [21, 22]. The structure of TelK (from phage φKO2), is shown in **Figure 6**; crystallized in a dimer conformation

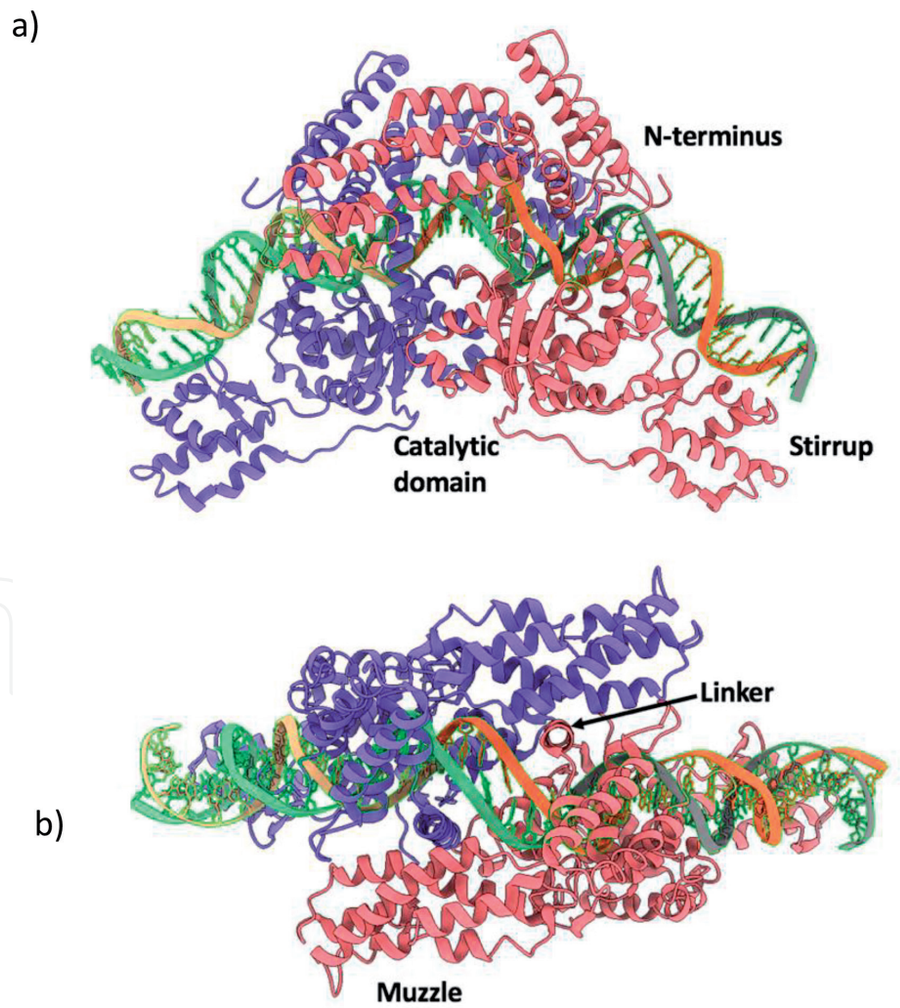


Figure 6.
Crystal structure of TelK dimer complexed with DNA. (a) Two monomers of TelK dimers at the recognition site and are held together by multiple transient protein-protein and protein-DNA interactions. The structure of each monomer is largely alpha helical, with mixed beta strands at the catalytic domain. (b) The same complex viewed from the N-terminus. The helical linker fits in the major groove of the DNA, it contacts the DNA on the opposite side to the rest of the protein. The structures presented were generated with using ChimeraX (Goddard et al., 2018) PDB ID: 2v6e [21].

complexed with the minimal cognate DNA sequence of 44-bp [PDB: 2V6E]. TelK has been divided into three core domains, all of which make contact with the DNA. These include, the muzzle at the N-terminus, the catalytic domain in the center, and the stirrup domain at the C-terminus. A long alpha-helical linker is also highlighted, and this connects the core catalytic and N-terminal domains.

One of the most striking observations of this structure is the level distortion: DNA is bent at roughly 73° parallel to the axis of symmetry [21]. This provides a valuable insight into the mechanism of telomere resolution by TelK and would appear to refute previous theories that the DNA is forced into a cruciform conformation [27]. Core substrate binding occurs at the N-terminus and the muzzle makes

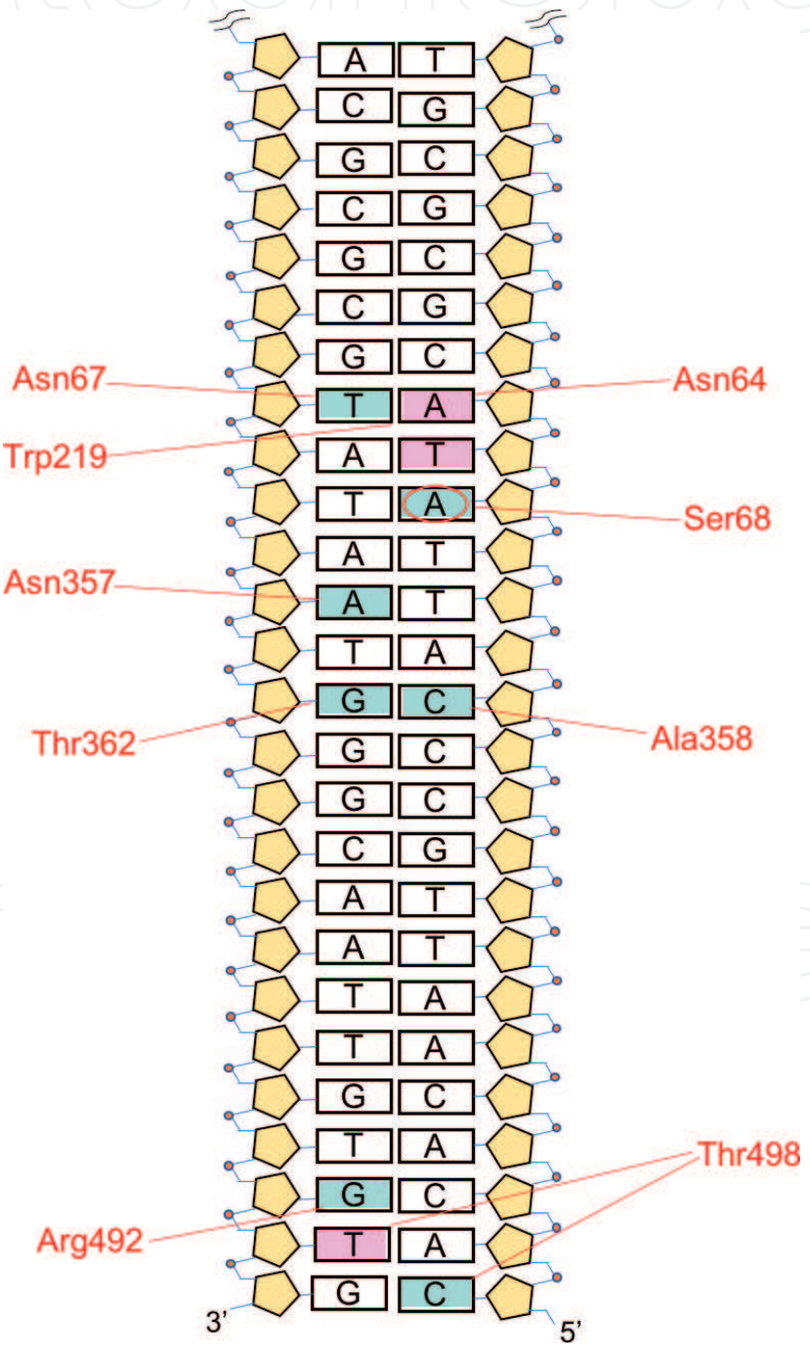


Figure 7.
TelK substrate recognition. One half of the TelK recognition sequence, nucleotides forming hydrogen bonds to protein are highlighted in cyan and those forming van der Waals interactions, in pink. Multiple amino acids interact with the DNA backbone and are not shown, those forming bonds with the bases are illustrated in red. Adenine at position 42 is circled, this residue forms hydrogen bonds with serine 68 of TelK and is therefore important for site recognition. Adapted from [21].

extensive contacts with the opposite subunit, strengthening the protein's structure and pushing the DNA into this strained conformation. The catalytic site is formed at the dimer interface, it binds to the opposite side of the substrate relative to the N-terminus and the helical linker that connects these two domains is fixed in the major groove. Extensive electrostatic interactions mediate the interaction.

Closer examination of the interactions between the DNA and protein reveal seven nucleotides that form hydrogen bonds with nearby residues (shown in cyan in **Figure** 7) and are presumed key for substrate recognition. This model is supported by the previously cited studies, whereby the natural PY54 substrate was effectively mutated into a sequence that could be processed by TelK. Adenine at position 42 is circled; this is one of the bases identified as forming hydrogen bonds with TelK and is one of the points that required mutating in order to form a cognate sequence. High salt has been shown to inhibit telomere resolution by protelomerases [16], this is possibly reflected by the extensive hydrophilic protein-DNA interactions that would be disrupted by an excess of ions.

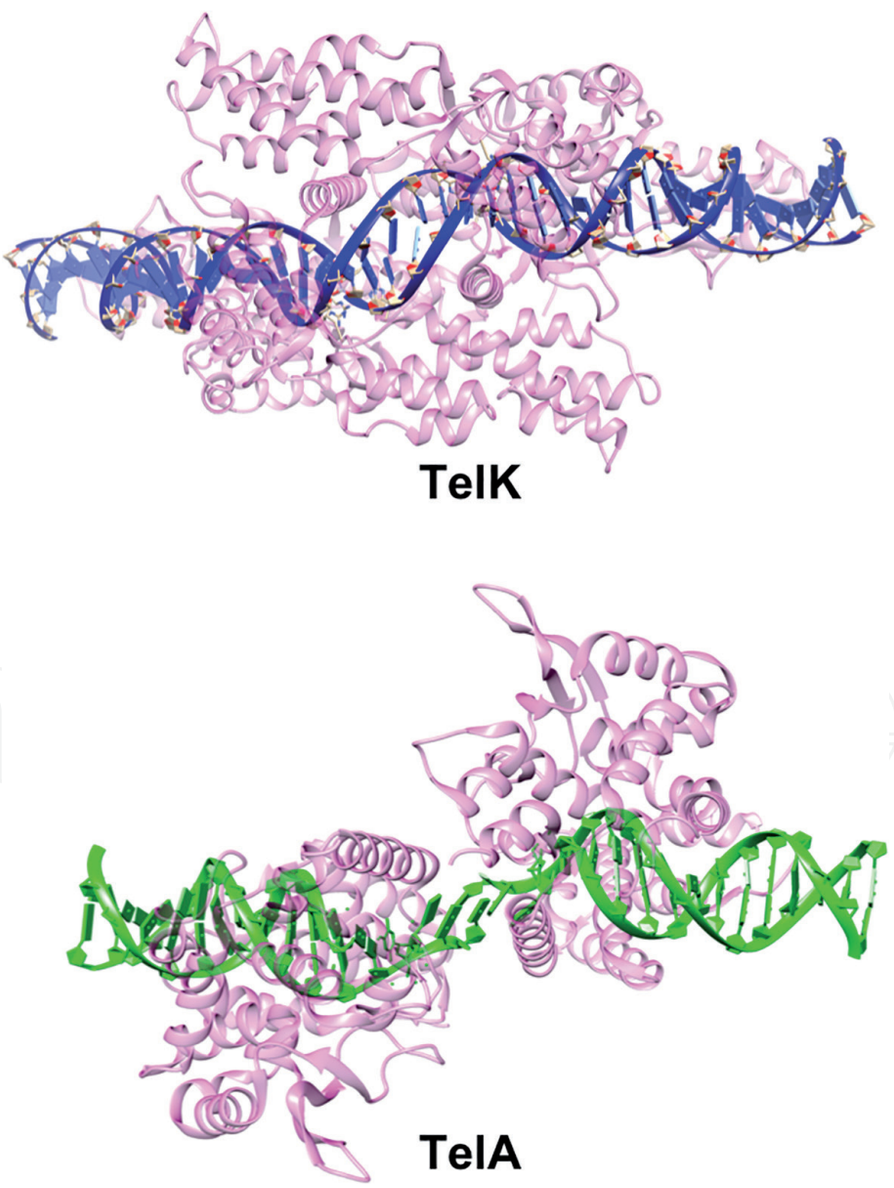


Figure 8.
Dimers of TelK and TelA complexed with DNA. The DNA is forced into the same disjointed structure down its helical axis [22]. The structures presented were generated with UCSF Chimera (Pettersen, at al., 2004) using PDB 2v6e (TelK) [21] and 4eop (TelA) [22].

The stirrup domain of TelK has a winged helix-turn-helix motif [21], it makes few contacts with the rest of the protein but extends the DNA binding interface. In stabilizing the strained substrate conformation, this part of the protein aids hairpin formation; however, it is nonessential for the cleavage reaction [21]. Following strand cleavage, the stored energy is released, and this drives dimer dissolution which is proceeded by spontaneous hairpin formation. The stirrup is not conserved among protelomerases and this provides further evidence to support the theory that the mechanism of telomere resolution varies between different systems.

The structure of the bacterial protelomerase, TelA has also been described via X-ray crystallography. TelA is considerably smaller than TelK, it lacks the stirrup and only consists of the catalytic and N-terminal domains. This design is similar to that of tyrosine recombinases, which are also typically composed of two domains [48]. The N-terminal 100 residues are poorly resolved in comparison to the rest of the protein, this area of low electron density suggests flexibility of the polypeptide chain. Comparing dimer substrate complexes of TelA to those of TelK (**Figure 8**) reveals a similar DNA conformation at the dimer interface. Extensive hydrogen bonds and van der Waals interactions are also involved in dictating the substrate specificity of TelA and the DNA exhibits the same disjunction down its helical axis [22].

6.2 Catalytic domain

The catalytic domain of TelK is a mixed alpha beta structure. **Figure 9** shows the core catalytic residues of TelK R275, K300, K380, R383 and H416. These act together to maintain catalytic activity and coordinate nucleophilic attack of the tyrosine. Side chains of the basic amino acids at positions one, three and four provide a hydrogen bonding network that coordinates the scissile phosphate and stabilizes the transition state [22]. In type IB topoisomerases, the pentad is usually composed of RKKRH/N [49], with basic residues 1 and 3 having the same stabilization effect. In these proteins, the second lysine residue has been shown to donate a proton to the

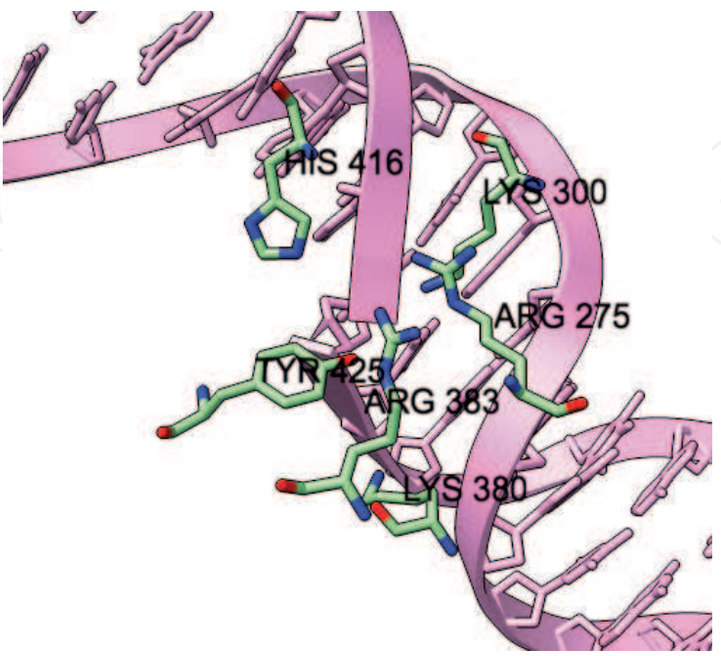


Figure 9.
Conformation of protelomerase active site residues. The residues of TelK, R275, K300, K380, R383 and H416 maintain the active site conformation and coordinate nucleophilic attack of the tyrosine. The structures presented were generated with ChimeraX (Goddard, T et al., 2018) using PDB 2v6e [21].

5′-OH leaving group, which aids its removal during cleavage [50]. It is feasible that the lysine in the protelomerase's active site, with its side chain positioned between the DNA O5' and nonbridging oxygen, also functions in this way and protonates the leaving group [21].

These crystal structures are invaluable when trying to decipher the mechanism of protelomerase-catalyzed telomere resolution. TelN and TelK have highly homologous sequences and can process the same target site [27], it is therefore likely the structures of these enzymes are analogous and information about TelN can be derived from examining the TelK structure. However, is important to note that the crystallized structure of TelK is not full length and lacks 100 residues from the C-terminus. This does not appear to significantly affect protein activity in vitro [21], but the significance in vivo is unconfirmed.

7. Evolutionary history

It is widely believed Tyrosine recombinases and type IB topoisomerases have arisen from a common ancestor [50]. These structurally and mechanically analogous proteins catalyze important DNA rearrangements, they use an active site tyrosine residue and form covalently bound protein-DNA intermediates. Similarities between this mechanism, and that of protelomerases, has led to the suggestion that these protein classes have a shared evolutionary relationship [51].

Key to the function of tyrosine recombinases is the conserved catalytic motif: "RKHRH" [52, 53]. Crystal structures and sequence alignments reveal that protelomerases share this catalytic pentad [27], with exception of the middle residue, which varies between systems. Deviation at this central point is also observed in tyrosine recombinases, where the central histidine may be replaced by an arginine, asparagine, lysine or tyrosine. In protelomerases; TelN and TelK both have a methionine at this position, PY54 contains a lysine, VHML has a histidine and the bacterial protelomerases of *B. burgdorferi* and *A. tumefaciens*, have a tyrosine. It has been found substituting the tyrosine to a histidine or lysine in TelA is completely tolerable and results in no loss of activity [7]. Whether these variations are of significance is yet to be determined.

Structural comparisons outside of the catalytic domain reveal low overall sequence homology at both the N- and C- terminus. ResT is smaller than the phage proteins, partial proteolysis separates the 449 amino acid protein into two domains [54], and sequence analysis predicts its architecture is more comparable to that typical of tyrosine recombinases. In addition, this protein has the unusual and surprising ability to synapse Holliday Junctions [51]. The reaction appears to be favored in conditions that are counterproductive to telomere resolution, such as negative supercoiling, or unsymmetrical *telR* sites [51]. This observation provides compelling evidence for the argument that ResT may have evolved from a recombinase [51]. Although the significance of Holliday Junction formation by ResT is still under investigation, it has been suggested this could be an obsolete ancestral property of the protein [4]. Additional evidence highlighting a relationship between tyrosine recombinases and protelomerases comes from the Flp recombinase and lambda integrase. It has been found that, under specific conditions, these enzymes have the ability to form hairpin products [55, 56]. Thus, indicating the relative ease with which recombinases may be converted to telomere resolving enzymes.

In ResT, additional catalytic residues outside of the core pentad are required for telomere resolution [57]. This indicates digression from a typical tyrosine recombinase type mechanism. Furthermore, mutation of ResT histidine 324, the fifth residue of the catalytic motif does not result in total loss of activity [57]. This would

mark an obvious disparity between these two classes of enzymes, if it was not for the observation that this residue is also not essential in Flp recombinases [58]. Here it appears the final amino acid has a structural rather than catalytic role [58], which could be mirrored in ResT.

Analysis of the DNA substrates reveal further fundamental differences between protelomerases and tyrosine recombinases. Aside from ResT, which it has been suggested can act on more than one target site, protelomerases are specific and have stringent substrate sequence requirements. The target site of one recombinase often includes many other related points, such as phage or bacterial attachment sites. Furthermore, the reaction is intramolecular and can require auxiliary proteins [59]. This implies a mechanism more involved and complex than that of telomere resolution. The reaction catalyzed by type IB topoisomerases is considerably simpler, as these proteins act as monomers [60] and do not display the same sequence specificity as protelomerases.

In conclusion, there are key differences between the catalytic mechanism of protelomerases and tyrosine recombinases/type IB topoisomerases. However, there are also significant similarities and whether these proteins have evolved from a common ancestor is difficult to determine. It is expected that the conversion of a tyrosine recombinase to an enzyme capable of telomere resolution would be accompanied by the linearization of plasmid DNA [4]. The data suggesting this could be achieved with relative ease adds support to the argument these proteins are related. Under certain conditions, recombinases can have topoisomerase activity, and topoisomerases can affect DNA strand exchanges [27]. This raises an interesting question as to whether protelomerases, under the correct conditions, may also be able to exhibit topoisomerase activity [27].

8. Applications in biotechnology

The unique properties of protelomerases, and the DNA structures they produce, makes them a valuable class of protein that have important applications in synthetic biology and biotechnology. Linear DNA does not exhibit the supercoiling associated with plasmids, as the ends are free to rotate. This makes them stable vectors invaluable for cloning difficult sequences. DNA that is rich in adenine and thymidine, or contains lots of short tandem repeats is typically hard, if not impossible, to clone into circular vectors [10]. A commercially available cloning vector, pJAZZ from Lucigen, is based on the linear N15 phage genome. pJAZZ is sold as part of a cloning system (BigEasy Kit), enabling the insertion of otherwise unclonable sequences into the vector for the creation of viable linear plasmids that can be transformed into cells. pJAZZ vectors encode RepA and TelN, essential for bidirectional replication and telomere resolution. Transcriptional terminators flank the cloning site, thus minimizing interference and preventing transcription into and out of this region. These modifications extend the cloning possibilities and allow for the insertion of large cDNAs or operons [10].

The pJAZZ system has been further modified to specifically enhance its efficiency for the production of in vitro transcribed (IVT) mRNA. IVT mRNA is a powerful therapeutic tool, enabling the transient expression of heterologous proteins [61, 62]. However, in order to optimize translation efficiency and mRNA stability, the poly(A) tail length of the mRNA needs to be defined and optimized [63]. In particular, mRNA with poly(A) tails >300 nucleotides and purines at the 3′ end have been demonstrated as highly effective and appear to exhibit improved translation properties compared to those with shorter poly(A) tracts [64]. Due to its linear structure, a pJAZZ derived plasmid, called p(Extended Variable Length)

(pEVL), can create poly(A) tracts of up to 500 bps in length. Furthermore, the residues at the 3′ end can be defined as either adenine or guanine. This has a significant advantage over conventional circular DNA, which cannot incorporate more than 174 bp of poly(A) tract without conferring extreme instability [64].

Mediphage Bioceuticals is a genetic medicine company that has also utilized the unique DNA processing capability of protelomerases for their technology. They have developed a one-step in vivo platform to produce linear covalently closed constructs, called ministring DNA [65]. These constructs are produced in *E. coli* cells that have been engineered to express the PY54 protelomerase under the control of a heat-inducible promoter. Following induction, the protelomerase processes precursor plasmids in the cell, in doing so it effectively separates the desired expression cassette from the bacterial plasmid backbone. The ministring DNA can then be purified and used as a vector for gene or cell therapies and gene editing. Although ministring production is reliant on large scale bacterial fermentation, the absence of bacterial DNA elements render these constructs preferential to plasmid DNA for medicinal applications. Furthermore, the constructs are typically smaller than plasmids, this enhances their transfection efficiency thus making them less toxic, as fewer transfection reagents are required [12]. They are also more resistant to the shear-induced degradation that large plasmids are highly susceptible to [66].

Another in vivo application of protelomerases has been explored by Katzen at colleagues, who used TelN to fragment an *E. coli* chromosome into smaller, autonomous units [67]. These proof of concept experiments were designed as a solution to the considerable difficulties associated with synthesizing and manipulating large, stable genetic elements. In splitting the chromosome into two smaller units, which together contain the essential components required for cell viability, they significantly simplified the problem. Not only are the smaller units of genetic material easier to manipulate, but each episome is of a size that it can be assembled without the need for an assembly host. This work may be extended to fragment and linearize other genomic elements of interest, in particular for the study of large units, >2 Mbp, which at present cannot be assembled and maintained in any biological platform [67].

Touchlight Genetics Ltd. has also utilized protelomerases for their technology. The platform they have developed is a purely in vitro DNA production process that eliminates all the major problems associated with using bacterial fermentation for DNA amplification. Their cell-free technology uses a phage DNA polymerase from Phi29 to produce large amounts of DNA concatemers from small amounts of starting plasmid DNA template. DNA concatemers are processed by the protelomerase TelN, to create linear covalently closed constructs that are marketed as dbDNA. The closed ended linear DNA construct is able to encode long, difficult DNA sequences which are not tolerated in high yield production of plasmid DNA due to selection pressures. The in vitro amplification technology produces DNA containing no bacterial origin of replication sequences or antibiotic selection marker. These vectors are capable of immunizing against influenza infection, with a response comparable to that of plasmid DNA [68], as well as improving tumor growth control in a Human papillomavirus (HPV) driven head and neck cancer model by delivering a therapeutic vaccine encoding for HPV16 E6 and E7 antigens [69]. Furthermore, dbDNA constructs have been used to generate functional lentiviral vectors [14] and are promising candidates for the production of recombinant adeno-associated virus (AAV) vectors [13]. This platform could therefore be important for gene therapy [70, 71] and this synthetic process for DNA production will have a broad range of application within the wider synthetic biology field.

9. Conclusion

Protelomerases are an interesting and unique class of protein. In forming telomeric structures at the ends of linear plasmids, they protect the genetic material from degradation and provide a novel solution to the end replication problem. They have also been implemented as an essential component of certain human pathogens and have important applications in synthetic biology. Despite this, protelomerases remain poorly characterised and many questions about their structure, function, mechanism and evolutionary history remain unanswered.

Much of our current knowledge has been obtained by combining information from crystal structures, analysing sequences and performing in vitro assays with protein and substrate variations. The results of these studies have enabled us to compare the protelomerases from different systems and it is clear that there is much variety within this protein family. In particular, ResT has been identified as having additional, largely unexplainable, functionality aside from telomere resolution. Potentially further characterization of the other protelomerases will lead to similar revelations.

Biochemical analysis of protelomerases from ΦHAP1, PY54 and VP58.2, will shed further light on the underlying properties that differentiate these telomere resolving enzymes. Such work could also explore the evolution of protelomerases and prokaryotes with linear plasmids. Similarities in the genome organization of telomere phage suggest a common ancestor [43] and whether the bacterial proteins originated from these is currently unknown. Introducing different phage into the same cell and determining their compatibility can give insight into evolutionary background [72]. An understanding of the relationship between phage lambda and telomere phage, in particular N15, will provide an interesting and important insight into how plasmid and phage can interact and evolve.

Crystal structures of TelK and TelA provide a solid starting point for research aiming to solve structures of protelomerases. These can also be used for homology modeling to infer information about proteins with high sequence similarities. Advances in the field of synthetic biology and protein engineering make detailed knowledge of these proteins even more valuable. An increased understanding of how they operate, and which parts are responsible for specific functionalities, will open up opportunities to produce variants with altered activities.

Appendices and nomenclature

dbDNA	doggybone DNA
bp	base pair
Tos	Telomerase occupancy site
Cos	cohesive end sites
Ori	origin of replication
IVT	in vitro transcribed
pEVL	p(Extended Variable Length)
AAV	adeno-associated virus
HPV	Human papillomavirus

Conflict of interest

SK has recently started a PhD jointly between Touchlight Genetics and Renos Savva; however, this has not relatively biased or affected the content of this chapter.

IntechOpen

IntechOpen

Author details

Sophie E. Knott, Sarah A. Milsom and Paul J. Rothwell*
Touchlight Genetics, Hampton, UK

*Address all correspondence to: paul.rothwell@touchlight.com

IntechOpen

© 2019 The Author(s). Licensee IntechOpen. This chapter is distributed under the terms of the Creative Commons Attribution License (http://creativecommons.org/licenses/by/3.0), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited. CC BY

References

[1] Ravin V, Ravin N, Casjens S, Ford ME, Hatfull GF, Hendrix RW. Genomic sequence and analysis of the atypical temperate bacteriophage N15. Journal of Molecular Biology. 2000;**299**(1):53-73

[2] Lan SF, Huang CH, Chang CH, Liao WC, Lin IH, Jian WN, et al. Characterization of a new plasmid-like prophage in a pandemic *Vibrio parahaemolyticus* O3:K6 strain. Applied and Environmental Microbiology. 2009;**75**(9):2659-2667

[3] Kobryn K, Chaconas G. ResT, a telomere resolvase encoded by the Lyme disease spirochete. Molecular Cell. 2002;**9**:195-201

[4] Chaconas G, Kobryn K. Structure, function, and evolution of linear replicons in *Borrelia*. Annual Review of Microbiology. 2010;**64**(1):185-202

[5] Kobryn K, Chaconas G. Hairpin telomere resolvases. Microbiology Spectrum. 2014;**2**(6):273-287

[6] Loh T, Elvitigala T, Wang C, Wollam A, Welsh EA, Liberton M, et al. The genome of Cyanothece 51142, a unicellular diazotrophic cyanobacterium important in the marine nitrogen cycle. Proceedings of the National Academy of Sciences of the United States of America. 2008;**75**(9):2659-2667

[7] Huang WM, DaGloria J, Fox H, Ruan Q, Tillou J, Shi K, et al. Linear chromosome-generating system of *Agrobacterium tumefaciens* C58: Protelomerase generates and protects hairpin ends. The Journal of Biological Chemistry. 2012;**287**(30):25551-25563

[8] Mir T, Huang SH, Kobryn K. The telomere resolvase of the Lyme disease spirochete, *Borrelia burgdorferi*, promotes DNA single-strand annealing and strand exchange. Proceedings of the National Academy of Sciences of the United States of America. 2013;**41**(22):10438-10448

[9] Huang SH, Cozart MR, Hart MA, Kobryn K. The *Borrelia burgdorferi* telomere resolvase, ResT, possesses ATP-dependent DNA unwinding activity. Nucleic Acids Research. 2017;**45**(3):1319-1329

[10] Godiska R, Mead D, Dhodda V, Wu C, Hochstein R, Karsi A, et al. Linear plasmid vector for cloning of repetitive or unstable sequences in *Escherichia coli*. Nucleic Acids Research. 2010;**38**(6):e88

[11] Godiska R, Ravin N, Mead D. Linear plasmid vector for cloning "Unclonable" DNA. BioTechniques. 2008;**45**(5):592

[12] Nafissi N, Slavcev R. Construction and characterization of an in-vivo linear covalently closed DNA vector production system. Microbial Cell Factories. 2012;**11**:154. DOI: 10.1186/1475-2859-11-154

[13] Karbowniczek K, Rothwell P, Extance J, Milsom S, Lukashchuk V, Bowes K, et al. Doggybone™ DNA: An advanced platform for AAV production. Cell & Gene Therapy Insights. 2017:731-738

[14] Karda R, Counsell JR, Karbowniczek K, Caproni LJ, Tite JP, Waddington SN. Production of lentiviral vectors using novel, enzymatically produced, linear DNA. Gene Therapy. 2019;**26**:86-92

[15] Tourand Y, Kobryn K, Chaconas G. Sequence-specific recognition but position-dependent cleavage of two distinct telomeres by the *Borrelia burgdorferi* telomere resolvase, ResT. Molecular Microbiology. 2003;**48**:901-911

[16] Deneke J, Ziegelin G, Lurz R, Lanka E. The protelomerase of temperate *Escherichia coli* phage N15 has cleaving-joining activity. Proceedings of the National Academy of Sciences of the United States of America. 2000;**97**(14):7721-7726

[17] Landry MP, Zou X, Wang L, Huang WM, Schulten K, Chemla YR. DNA target sequence identification mechanism for dimer-active protein complexes. Nucleic Acids Research. 2013;**41**(4):2416-2427

[18] Ravin NV. Mechanisms of replication and telomere resolution of the linear plasmid prophage N15. FEMS Microbiology Letters. 2003;**221**(1):1-6

[19] Lucyshyn D, Huang SH, Kobryn K. Spring loading a pre-cleavage intermediate for hairpin telomere formation. Nucleic Acids Research. 2015;**43**(12):6062-6074

[20] Chou S, Chin K, Wang AH. Unusual DNA duplex and hairpin motifs. Nucleic Acids Research. 2003;**31**(10):2461-2474

[21] Aihara H, Huang WM, Ellenberger T. An interlocked dimer of the protelomerase TelK distorts DNA structure for the formation of hairpin telomeres. Molecular Cell. 2007;**27**(6):901-913

[22] Shi K, Huang WM, Aihara H. An enzyme-catalyzed multistep DNA refolding mechanism in hairpin telomere formation. PLoS Biology. 2013;**11**(1):e1001472

[23] Bankhead T, Chaconas G. Mixing active-site components: A recipe for the unique enzymatic activity of a telomere resolvase. Proceedings of the National Academy of Sciences. 2004;**101**(38):13768-13773

[24] Briffotaux J, Kobryn K. Preventing broken *Borrelia* telomeres ResT couples dual hairpin telomere formation with product release. The Journal of Biological Chemistry. 2010;**285**(52):41010-41018

[25] Casjens S, Murphy M, Sampson L. Telomeres of the linear chromosomes of Lyme disease spirochaetes: Nucleotide sequence and possible exchange with linear plasmid telomeres. Molecular Microbiology. 1997;**26**:581-596

[26] Deneke J, Lurz R, Lanka E. Phage N15 telomere resolution. The Journal of Biological Chemistry. 2002;**277**(12):10410-10419

[27] Huang WM, Joss L, Hsieh T, Casjens S. Protelomerase uses a topoisomerase IB/Y-recombinase type mechanism to generate DNA hairpin ends. Journal of Molecular Biology. 2004;**337**(1):77-92

[28] Kobryn K, Burgin AB, Chaconas G. Uncoupling the chemical steps of telomere resolution by ResT. The Journal of Biological Chemistry. 2005;**280**(29):26788-26795

[29] Zabala B, Hammerl JA, Espejo RT, Hertwig S. The linear plasmid prophage Vp58. 5 of *Vibrio parahaemolyticus* is closely related to the integrating phage VHML and constitutes a new incompatibility group of telomere phages. Journal of Virology. 2009;**83**(18):9313-9320

[30] Hertwig S, Klein I, Lurz R, Lanka E, Appel B. PY54, a linear plasmid prophage of *Yersinia enterocolitica* with covalently closed ends. Molecular Microbiology. 2003;**48**(4):989-1003

[31] Mobberley JM, Authement RN, Segall AM, Paul JH. The temperate marine phage HAP-1 of *Halomonas aquamarina* possesses a linear plasmid-like prophage genome. Journal of Virology. 2008;**82**(13):6618-6630

[32] Black L. DNA packaging in dsDNA. Annual Review of Microbiology. 1989;**43**:267-292

[33] Casjens S. Evolution of the linear DNA replicons of the *Borrelia* spirochetes. Molecular Microbiology. 1999;**2**(5):529-534

[34] Ravin NV. N15: The linear phage-plasmid. Plasmid. 2011;**65**(2):102-109

[35] Mardanov AV, Ravin NV. Conversion of linear DNA with hairpin telomeres into a circular molecule in the course of phage N15 lytic replication. Journal of Molecular Biology. 2009;**391**(2):261-268

[36] Ziegelin G, Scherzinger E, Lurz R, Lanka E. Phage P4 alpha protein is multifunctional with origin recognition, helicase and primase activities. The EMBO Journal. 1993;**12**(9):3703-3708

[37] Orejas RD, Ziegelin G, Lurz R, Lanka E, Genetik MM, Schuster A. Phage P4 DNA replication in vitro. Nucleic Acids Research. 1994;**22**(11):2065-2070

[38] Rybchin VN, Svarchevsky AN. The plasmid prophage N15: A linear DNA with covalently closed ends. Molecular Microbiology. 1999;**33**:895-903

[39] Landschulz WH, Johnson PF, Mcknight SL, Landschulz WH, Johnson PF, Mcknight SL. Structure common to a new class of the leucine zipper: A hypothetical DNA binding proteins. Science. 1988;**240**(4860):1759-1764

[40] Ravin NV, Strakhova TS, Kuprianov V. The protelomerase of the phage-plasmid N15 is responsible for its maintenance in linear form. Journal of Molecular Biology. 2001;**312**(5):899-906

[41] Feiss M, Young J, Sultana S, Patel P, Sippy JDNA. Packaging specificity of bacteriophage N15 with an excursion into the genetics of a cohesive end mismatch. PLoS ONE. 2015;**10**(12):e0141934

[42] Papagiannis CV, Sam MD, Abbani MA, Yoo D, Cascio D, Clubb RT, et al. Fis targets assembly of the Xis nucleoprotein filament to promote excisive recombination by phage lambda. Journal of Molecular Biology. 2007;**367**(2):328-343

[43] Replication RNV. Maintenance of linear phage-plasmid N15. Microbiology Spectrum. 2015;**3**(1):1-12

[44] Baril C, Richaud C, Baranton G. Linear chromosome of *Borrelia burgdorferi*. Research in Microbiology. 1989;**140**(8):507-516

[45] Picardeau M, Lobry JR, Joseph B, Bernard C. Physical mapping of an origin of bidirectional replication at the centre of the *Borrelia burgdorferi* linear chromosome. Molecular Microbiology. 1999;**32**(2):437-445

[46] Chaconas G, Stewart PE, Tilly K, Bono JL, Rosa P. Telomere resolution in the Lyme disease spirochete. The EMBO Journal. 2001;**20**(12):3229-3237

[47] Byram R, Stewart PE, Rosa P. The essential nature of the ubiquitous 26-kilobase circular replicon of *Borrelia burgdorferi*. Journal of Bacteriology. 2004;**186**(11):3561-3569

[48] Yang W, Mizuuchi K. Site-specific recombination in plane view. Structure. 1997;**5**(11):1401-1406

[49] Cheng C, Kussie P, Pavletich N, Shuman S. Conservation of structure and mechanism between eukaryotic topoisomerase I and site-specific recombinases. Cell. 1998;**92**:841-850

[50] Krogh BO, Shuman S. A poxvirus-like type IB topoisomerase family in bacteria. Proceedings of the National Academy of Sciences. 2002;**99**(4):1853-1858

[51] Kobryn K, Briffotaux J, Karpov V. Holliday junction formation by the *Borrelia burgdorferi* telomere resolvase, ResT: Implications for the origin of genome linearity. Molecular Microbiology. 2009;**71**(5):1117-1130

[52] Grainge I, Jayaram M. The integrase family of recombinases: Organization and function of the active site. Molecular Microbiology. 1999;**33**:449-456

[53] Esposito D, Scocca JJ. The integrase family of tyrosine recombinases: Evolution of a conserved active site domain. Nucelic Acids Research. 1997;**25**(18):3605-3614

[54] Tourand Y, Lee L, Chaconas G. Telomere resolution by *Borrelia burgdorferi* ResT through the collaborative efforts of tethered DNA binding domains. Molecular Microbiology. 2007;**64**:580-590

[55] Lee J, Tonozuka T, Jayaram M. Mechanism of active site exclusion in a site-specific recombinase: Role of the DNA substrate in conferring half-of-the-sites activity. Genes & Development. 1997;**11**:3061-3071

[56] Nash HA, Robertson CA. Heterduplex substrates for bacteriophage lambda site-specific recombination: Cleavage and strand transfer products. The EMBO Journal. 1989;**8**(11):3523-3533

[57] Deneke J, Burgin AB, Wilson SL, Chaconas G. Catalytic residues of the telomere resolvase ResT. The Journal of Biological Chemistry. 2004;**279**(51):53699-53706

[58] Chen Y, Rice PA. The role of the conserved Trp 330 in Flp-mediated recombination. The Journal of Biological Chemistry. 2003;**278**(27):24800-24807

[59] Nunes-düby SE, Kwon HJ, Tirumalai RS, Ellenberger T, Landy A. Similarities and differences among 105 members of the Int family of site-specific recombinases. Nucleic Acids Research. 1998;**26**(2):391-406

[60] Wang JC. DNA topoisomerases. Annual Review of Biochemistry. 1996;**65**:635-692

[61] Kormann MSD, Hasenpusch G, Aneja MK, Nica G, Flemmer AW, Herber-jonat S, et al. Expression of therapeutic proteins after delivery of chemically modified mRNA in mice. Nature Biotechnology. 2011;**29**(2):154-157

[62] Bangel-ruland N, Fernández EF, Leier G, Leciejewski B, Rudolph C, Rosenecker J, et al. Cystic fibrosis transmembrane conductance regulator-mRNA delivery: A novel alternative for cystic fibrosis gene therapy. The Journal of Gene Medicine. 2013;**15**:414-426

[63] Gallie D. The cap and poly (A) tail function synergistically to regulate mRNA translational efficiency. Genes & Development. 1991;**5**:2108-2116

[64] Grier AE, Burleigh S, Sahni J, Clough CA, Cardot V, Choe DC, et al. pEVL: A linear plasmid for generating mRNA IVT templates with extended encoded poly(A) sequences. Molecular Therapy--Nucleic Acids. 2016;**5**(4):e306

[65] Wong S, Lam P, Nafissi N, Denniss S, Slavcev R. Production of double-stranded DNA ministrings. Journal of Visualized Experiments. 2016;**108**:3-9

[66] Catanese DJ, Fogg JM, Schrock DE, Gilbert BE, Zechiedrich L. Supercoiled minivector DNA resists shear forces associated with gene therapy delivery. Gene Therapy. 2011;**19**(1):94-100

[67] Liang X, Baek C, Katzen F. *Escherichia coli* with two linear chromosomes. ACS Synthetic Biology. 2013;**2**(12):734-740

[68] Scott VL, Patel A, Villarreal DO, Hensley SE, Ragwan E, Yan J, et al. Novel synthetic plasmid and doggybone DNA vaccines induce neutralizing antibodies and provide protection from lethal influenza challenge in mice. Human Vaccines & Immunotherapeutics. 2015;**11**(8):1972-1982

[69] Allen A, Wang C, Caproni LJ, Sugiyarto G, Harden E, Douglas LR, et al. Linear doggybone DNA vaccine induces similar immunological responses to conventional plasmid DNA independently of immune recognition by TLR9 in a pre-clinical model. Cancer Immunology, Immunotherapy. 2018;**67**(4):627-638

[70] Daya S, Berns KI. Gene therapy using adeno-associated virus vectors. Clinical Microbiology Reviews. 2008;**21**(4):583-593

[71] Naso MF, Tomkowicz B, Perry WL, Strogl WR. Adeno-associated virus (AAV) as a vector for gene therapy. BioDrugs. 2017;**31**(4):315-332

[72] Hammerl JA, Klein I, Appel B, Hertwig S. Interplay between the temperate phages PY54 and N15, linear plasmid prophages with covalently closed ends. Journal of Bacteriology. 2007;**189**(22):8366-8370