Mechanisms of RecA filament nucleation on ssDNA by the DprA protein

Irina Bakhlanova[1], Aleksandr Alekseev[2], Maria Yakunina[2], Natalia Morozova[2], Mikhail Khodorkovskii[2], Michael Petukhov[1], Dmitry Baitin[2]

Affiliation

[1] Petersburg Nuclear Physics Institute Named by B.P. Konstantinov of National Research Center «Kurchatov Institute», Gatchina, Russian Federation.

[2] Peter the Great St. Petersburg Polytechnic University, St. Petersburg, Russian Federation

Abstract

The DprA protein has been implicated in the protection of incoming DNA. However, the presence of dprA gene family members, also known as smf, across diverse bacterial species suggests a broader functionality for their gene products. We examined the role of the *Escherichia coli* DprA/Smf homologue in conjugation. Remarkably, *Bacillus subtilis* dprA/smf can complement an E. coli dprA mutant, indicating interchangeability of dprA/smf genes between competent and non-competent species in conjugational processes. The DprA protein forms a complex with DNA, facilitating the nucleation of RecA protein filaments onto circular single-stranded DNA coated with SSB protein. To focus on RecA nucleation, we employed short DNA oligonucleotides that restrict RecA-DNA binding but allow for DprA-RecA-DNA binding. Analysis of dATPase activity revealed that RecA-DNA complexes were readily formed only with olig50, while DprA-RecA-DNA complexes were also feasible with olig21. Combining experimental data with a full-atomic model of the RecA-DprA-ssDNA complex's spatial structure, we proposed a molecular mechanism for DprA-mediated loading of RecA proteins onto ssDNA. Our findings suggest that only one DprA-ssDNA interaction can occur sterically, occupying one strand of ssDNA in the complex.

Keywords: RecA, conjugation, DprA,

Abbreviations:
FRE frequency of recombination exchanges per DNA unit length
SSB single-stranded DNA binding protein
ssDNA single-strand DNA
poly(dT) extended polymers of deoxythymidine
nt nucleotide(s)
EMSA electrophoretic mobility shift assay

Introduction

DprA/Smf is a widely distributed protein that initiates natural transformation, and its initiation activity is constrained by the DNA uptake and genetic recombination machinery. DprA is present in both transformable and noncompetent species, indicating its potential multifunctionality [1-4]. The ubiquitous presence of the DprA domain across bacterial species, coupled with a high degree of sequence conservation, suggests its vital functional role. DprA's role as a transformation factor in numerous Gram-positive species has been well-established, with extensive studies on DprA proteins from B. subtilis, D. radiodurans, and S. pneumoniae [5-

8]. However, the primary function of DprA in non-transforming Gram-negative bacteria like E. coli remains unclear. DprA/Smf binds to environmental DNA, playing a crucial role in RecA-independent plasmid transformation or RecA-dependent chromosomal transformation. Its concerted binding to dsDNA leads to the processing of one of the strands, facilitating the uptake of internalized ssDNA to replace homologous chromosomal sequences. The ability of DprA from various bacterial species to bind both ssDNA and dsDNA has been demonstrated [5, 7, 9]. Additionally, the recruitment of RecA protein by DprA/Smf mediates the partial displacement of the SSB protein from ssDNA [5, 10].

DprA has been identified in many bacterial species not known to undergo transformation, leading to the hypothesis that DprA functions beyond transformation in these bacteria. The E. coli *dprA* gene, called *smf*, was shown not to play a role in transformation [11]. However, heterospecific interaction between RecA and DprA has been observed. The *smf* from *E. coli* has the ability to restore natural transformation in *Haemophilus influenza* [11]. This suggests an evolutionarily conserved DprA (Smf)-RecA interaction, highlighting its biological significance. The *E. coli* DprA/Smf is a 370-amino acid residue protein with a molecular weight of 42 kDa. While the crystal structure of full-length E. coli DprA/Smf is unavailable, the structure of *S. pneumoniae* DprA has been elucidated. The *S. pneumoniae* DprA structure consists of the association of a sterile alpha motif domain and a Rossmann fold, forming tail-to-tail dimers. Key interaction residues positioning on *the S. pneumoniae* DprA structure indicates an overlap of DprA–DprA and DprA–RecA interaction surfaces. Overall, the structural features of DprA protein suggest a competitive interaction between DprA and RecA for the same binding surface motif, both being equally essential for transformation proficiency [6].

To determine whether *E. coli* DprA/Smf is capable of functioning during conjugation, we investigated the role of DprA/Smf protein when expressed from a plasmid. By controlling the time of protein induction and temperature incubation, we demonstrate an almost twenty-fold increase in FRE. Additionally, we tested the effect of *B. subtilis* dprA on conjugation for comparative purposes, as it is well-characterized and shares a significant identity and similarity (42% and 62%, respectively) with E. coli DprA/Smf.

There are no published accounts that the *E. coli* DprA protein has been purified for biochemical studies, although not for a lack of trying. The *E. coli* Dpra is highly susceptible to proteolytic breakdown and is generally insoluble in aqueous solutions. In the current study, we demonstrate the effect of the *B. subtilis* DprA protein on E.coli RecA-mediated dislodging of SSB protein and DNA oligonucleotides binding using a biochemical system for ATP hydrolysis

measuring. Here we report a significant stimulation of the RecA nucleation step of DNA binding by the DprA protein. Firstly, DNA oligonucleotide was used as the shortest substrate to experimental measure DNA site size of DprA-RecA complex. In vitro data revealed the capacity of just one DprA-ssDNA connection that provides a suitable platform to promote the loading of RecA on that shortest ssDNA substrate. We also propose a full-atomic model of the spatial structure of the RecA–DprA–ssDNA complex build using molecular modeling. The model describes sterically possible form of the RecA–DprA–ssDNA complex capable to explain the experimental data.

Result

DprA significantly increase the frequency of recombinational exchanges

Constant presentation of new ssDNA occurs during conjugation, as the donor DNA is transferred into recipients. The incoming donor ssDNA undergoes conversion to a duplex form through the synthesis of Okazaki fragments, forming gapped DNA [12]. Throughout mating, the donor Hfr KL227 transfers markers into recipients sequentially in the order of leu+, ara+, and thr+. The recombination exchanges in this region of the E. coli map are adequately described mathematically by the Haldane formula. Over the past 20 years, a method for measuring recombination exchanges to assess elevated or depressed levels of recombination in vivo has been developed. This method has the advantage of detecting essentially all types of genetic exchanges that might occur during conjugation [13].

At least two quantitative parameters have been utilized to characterize conjugational recombination - the yield of recombinants and the linkage of two transferred donor markers. The frequency of recombinational exchanges is expressed as FRE, the average number of exchanges per one E. coli genome equivalent (100 minutes). For wild-type E. coli, FRE = 5.0. The FRE approach aligns with methodologies developed by others [14]. We explored a potential role for smf in conjugation through Hfr conjugation experiments, utilizing both DprA/Smf from E. coli and B. subtilis. The genes encoding the E. coli DprA/Smf protein or the B. subtilis DprA protein were expressed in the presence of IPTG. Both DprA proteins exhibited a low level of recombinant yield. Surprisingly, B. subtilis DprA not only acted as an inhibitor of the yield of recombinants but proved to be a much more potent inhibitor than DprA/Smf from E. coli. Previous findings showed that some hyper-rec events observed for RecA mutants led to a decrease in the yield of recombinants. This suggests that the inhibitory effect of DprA on the yield of recombinants may correspond to its toxic effect on DNA intermediates [15].

The differential effects of the two DprA proteins are readily observed in FRE values under normal conditions. There was a 2.6 and 6.4-fold increase in FRE values, relative to rec+, in the transconjugants observed in cells expressing E. coli DprA/Smf or B. subtilis DprA. The data presented in this section indicate that, compared to E. coli DprA/Smf, the B. subtilis DprA protein appears to be a stronger regulator of HR activity, measured through FRE, the stimulatory role of which was measured at 37 °C. To facilitate the accumulation of a database that is more optimally matched for protein activity, we conducted the following studies at a 30 °C temperature of protein induction. Surprisingly, the FRE level increased for both proteins. In effect, the 30 °C condition of protein induction led to a 17 – 19 fold increase in FRE values. The DprA proteins contributed equally to the recombination potential under conditions of protein over-expression at 30 °C.

As previously described, normal FRE values for the wild-type background appear to be induced in a temperature-dependent manner, showing a 2-fold increase in FRE at 42°, respectively [13]. This suggests that temperatures as high as 37 °C inhibit the intracellular amount or activity of proteins or both. This result leads to two major conclusions. The primary conclusion is that the DprA protein can be viewed as an important regulator of HR in Gram-negative bacterial conjugation with a strong stimulatory effect. Second, DprA from B. subtilis efficiently replaces DprA from E. coli in E. coli cells, showing a similar FRE value, regardless of the mechanism of action behind the positive effect.

Table 1a The dependence of the FRE value on the increased expression of DprA in transconjugants of the AB1157 line crossed with the KL227 donor, at 37 degrees.

	Induction time and T^0C	Linkage (*thr*$^+$-*leu*$^+$)	Yield of Thr+Strr recombinants (% to donors)	FRE	Δ FRE
ABpT7	1 h	0.930±0.025 (600)	1.6%	4.5±1.0	1.0
ABpT7 pDprA$_{Ec}$	37^0C	0.836±0.032 (600)	0.71%	11.7±1.5	2.6
ABpT7 pDprABs		0.689±0.021 (400)	0.22%	28.9±5.3	6.4

Table 1b The dependence of the FRE value on the increased expression of DprA in transconjugants of the AB1157 line crossed with the KL227 donor, at 30 degrees.

	Induction time and T^0C	Linkage (*thr*$^+$-*leu*$^+$)	Yield of Thr+Strr recombinants (% to donors)	FRE	Δ FRE
ABpT7	1 h	0.949±0.024 (600) *	1.7%	3.2±0.6	1.0
ABΔDprA	30^0C	0.943±0.018 (600)	3.1%	3.6±0.5	1.1
ABpT7 pDprA$_{Ec}$		0.588±0.025 (800)	0.2%	55.5±6.3	17.3
ABpT7 pDprABs		0.564±0.036 (600)	0.02%	61.2±8.5	19.1

*numbers in parentheses are the numbers of clones analyzed

Deletion of Smf confers a growth disadvantage

DNA metabolism comprises seemingly distinct processes tightly interlinked. The genome must be protected, replicated, expressed, organized, and segregated. All processes of DNA metabolism must share the same chromosomal substrate. Recombination systems ensure DNA replication success in all cells, facilitating the reconstitution of the replication forks that have collapsed upon encounters with those transient template strand discontinuities. The roles of the DprA protein in replication and cell division are not apparent. However, indirect evidence that dprA may be involved in cell division has been provided by comparative proteomic analysis of N. meningitides [16]. Previously, no differences in growth rate or viability were observed between the wild-type cells and the smf mutant in E. coli [11]. To explore the deleterious consequences of smf deletion, we first examined cellular growth rates. The smf deletion in E. coli results in about a 2-fold reduction in cell growth rates. The difference appears moderate, but it proved highly reproducible over the course of 4 trials. For the growth experiment with E. coli, we used a different medium. As shown in Fig 1., ΔdprA cells grow somewhat more slowly in minimal medium (MM) than cells grow in Luria-Bertani (LB) medium. Thus, our current work on DprA's role in the growth rate suggests that its function could be used in chromosome replication or cell division.

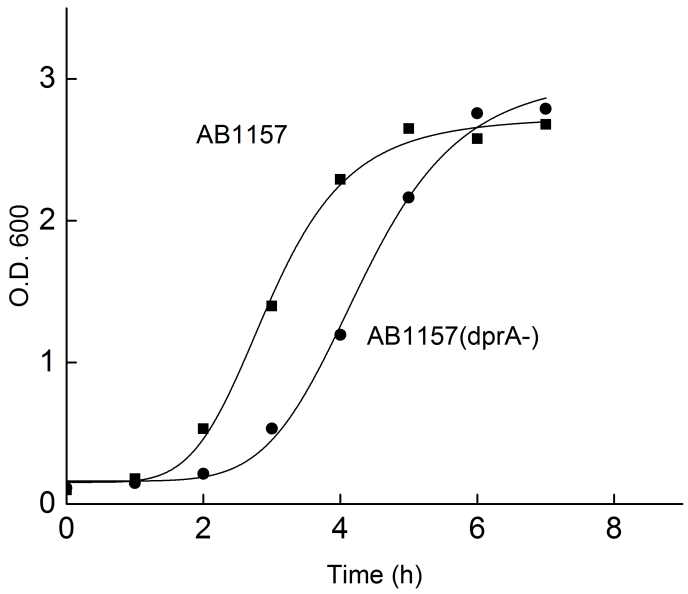

Fig.1 Growth disadvantage conferred by *ΔdprA/smf*. The growth curve of the *smf* mutant is compared to the wild type.

DprA promotes SSB dislodging from ssDNA

Previously, the ability of The B.subtilis RecA protein to displace barriers such as the B.subtilis SSB protein on single-stranded phage circular DNA was demonstrated to be enhanced in the presence of B. subtilis DprA protein. [7, 17]. The quantitative evaluation of this reaction is based on alterations in the RecA-dependent ATP hydrolysis rate. It's essential to note that the displacement process appears to be multifaceted. B.subtilis DprA protein interacts with B.subtilis SSB protein, creating nucleation sites for B.subtilis RecA protein on single-stranded DNA. Concurrently, B.subtilis DprA protein directly engages with B.subtilis RecA protein, increasing its affinity for single-stranded DNA.

Given that, in our experimental setup, RecA and SSB proteins are from E.coli, and only DprA protein is from B.subtilis, we aimed to determine whether this effect would be confirmed in our in vitro system. We replicated the experimental conditions similar to those used for the homogenous protein system from B.subtilis [17], i.e., an excess of phage DNA (12 μM) over E.coli RecA (1 μM). Simultaneously, the concentration of SSB protein was sufficient to bind all DNA present in the reaction - 0.5 μM of protein. Prior to the addition of RecA, DNA was

preincubated with SSB and subsequently with DprA proteins for the specified time interval. The reaction exhibited active displacement of SSB protein, evidenced by an increase in ATPase with a rise in the concentration of the regulatory protein DprA (Figure 2a).

The result suggests that Bacillus DprA effectively stimulates E.coli RecA under analogous conditions and concentration ratios as observed in the referenced study. In the initial minutes, a lag phase of ATP hydrolysis was observed, indicating a temporary delay in filament nucleation due to the obstacle presented by SSB. Subsequently, the ATP hydrolysis kinetics reached a steady-state, indirectly reflecting the proportion of each participating protein remaining bound to DNA. With a DprA protein concentration above 0.1 μM, the stimulating effect, ceased to increase after saturation was reached. Hence, the impact of heterospecific protein interactions, considering the provided concentrations, aligns with data on homospecific interactions published earlier. Ultimately, these findings affirm the existence of heterospecific interactions between E.coli and B.subtilis proteins in vitro, similar to what has been established for various strains previously, signifying the evolutionary conservation of protein-protein interactions. In this context, the identification of heterospecific interactions between DprA and RecA prompted us to employ RecA, available in its pure form, for a more comprehensive exploration of the impact of DprA on the interaction of RecA with ssDNA in vitro.

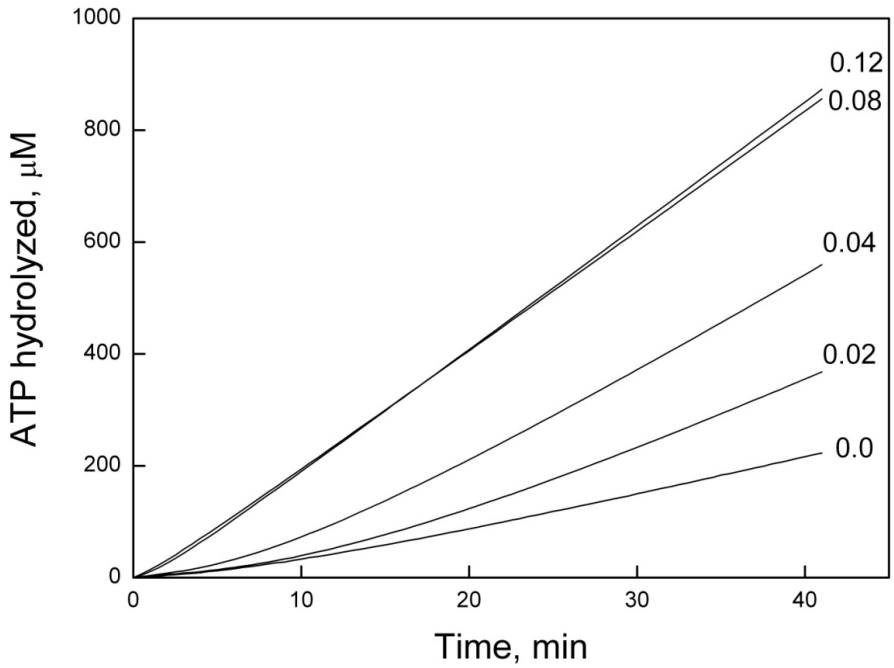

Fig. 2a *DprA protein antagonises SSB function. (a) The protective effects of SSB are not species-specific. The reactions were carried out as described under "Experimental Procedures" and contained 1 μm RecA protein, 12 μm M13mp8 ssDNA. The DprA protein concentration was varied as indicated.*

The quantitative equilibrium between SSB and RecA proteins at a steady state during their competitive interplay is affected not only by their concentrations and order in which they are added to the reaction mixture, but also by the nuances of the DNA architecture. If, on a circular DNA of any length, RecA tends to displace SSB, then on linear molecules with a free 5' end, conversely, SSB displaces RecA, as the disassembly of the RecA filament progresses in the 5'-3' direction [18, 19]. As the filaments dissociate from the ends, the time of disassembly depends on the number of ends. The shorter the linear DNA molecule, the shorter filament would dissociate in a shorter time and hence promptly replaced by the SSB protein. The kinetic outcomes of RecA filament disassembly are predictably different when long ssDNAs are substituted for poly(dT). Natural DNA molecules, such as long phage DNA, exhibit diverse secondary structures hindering filament polymerization. Known for unraveling DNA hairpins, the SSB protein aids the filament in extending along the DNA.

The B. subtilis DprA protein, as documented [7], stimulates the annealing of complementary strands, even in the presence of SSB. Consequently, each of these regulatory proteins may indirectly influence filamentation efficiency through DNA structure, adding complexity to data interpretation. In this particular experiment, a poly(T) oligonucleotide was used to eliminate potential side effects associated with circular single-stranded DNA, focusing on the direct effects of protein-protein interactions involving DprA, SSB, and RecA. The RecA concentration this experiment is ample, even in excess, to bind the entire poly(dT) oligonucleotide. As illustrated, SSB promptly, almost instantaneously, displaces RecA from the poly(dT), consistent with previously known data (Figure 2b). However, the DprA protein completely hinders filament disassembly, as the ATP hydrolysis rate in its presence remains unchanged following the addition of SSB. The observed steady state ATP hydrolysis in the presence of DprA protein remains constant, unaffected by the order of protein addition. Even if the poly(dT) is already bound by SSB protein, the kinetics are reversible. The newly established steady-state ATP hydrolysis is indistinguishable from the initial one. Given the poly(dT) oligonucleotide's lack of secondary structure (hairpins), all effects can be attributed to direct protein-protein interactions. Within this experiment, the sequence of SSB protein replacement by other proteins remains ambiguous. One mechanism was that DprA initiates binding first, displacing SSB protein by interaction and providing an available platform on DNA for RecA assembly [7, 20]. Another possible mechanism entails RecA and DprA engage with DNA simultaneously within a consolidated complex.

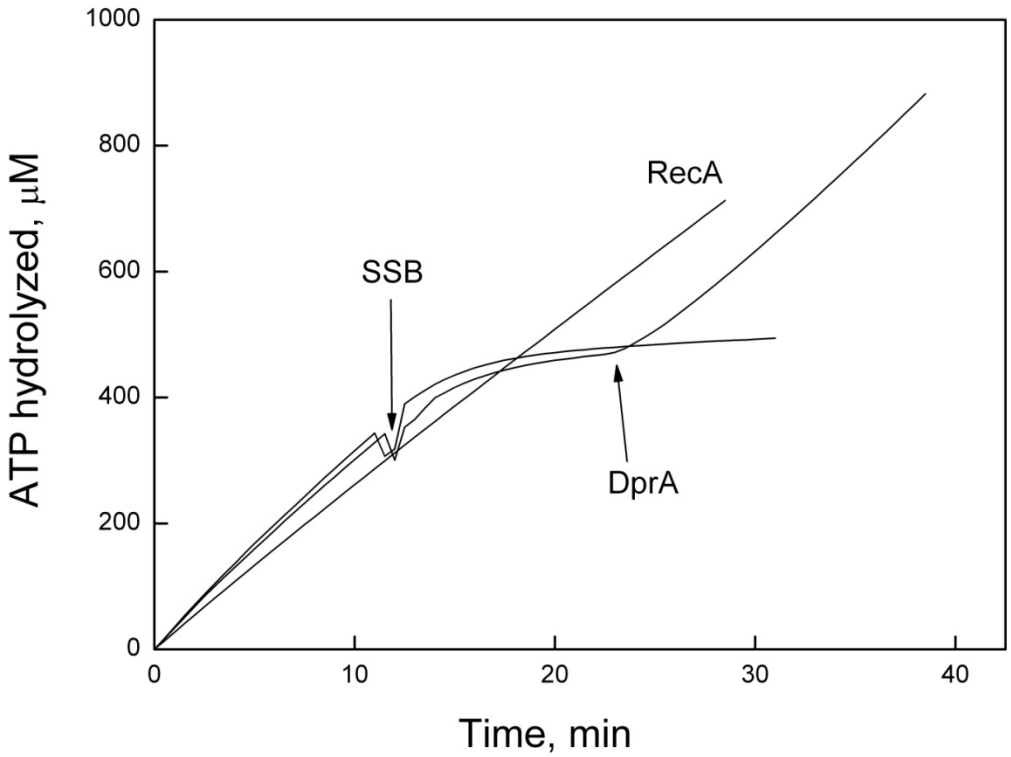

Fig. 2b (b) DprA can reverse the effects of SSB protein. *The reactions were carried out as described under "Experimental Procedures" and contained 3 μm RecA protein, 5 μm poly(dT).*

These two mechanisms are not mutually exclusive. The second mechanism, which ascribes greater significance to the interaction between RecA and DprA, is implied by the considerably enhanced capability of DprA to initiate RecA filamentation on the DNA.

DprA facilitate RecA nucleation onto short oligonucleotides

To examine the effect of DprA on RecA's nucleation ability, RecA was pre-incubated with oligos as short as 34 nucleotides, followed by the addition of DprA. As previously documented, the affinity of the RecA filament for the DNA substrate increased with the length of the ssDNA oligo, resulting in the formation of a highly stable DNA–RecA complex with oligo lengths exceeding 50 nucleotides [21, 22]. Varying the length of the substrate helps to isolate the nucleation phase from filament extension. Given the cooperative nature of monomer binding, an insufficient number of subunits in the filament compromises the complex's stability. The other current view is that the protein nucleates as an oligomer of 5 subunits or even as a dimer, forming a compact cluster on the DNA. [23, 24]. Anyway, in fact, a short DNA oligonucleotide with a length of 30–40 nucleotides proves to be a poor substrate for filament formation. The

interaction of the protein with DNA of such limited length is accompanied by only modest ATP hydrolysis. However, as the DNA length increases, the RecA-mediated hydrolysis rate gradually rises, reaching its peak values. A DNA oligonucleotide of 34 nucleotides can accommodate only 10-11 monomers. Considering some dynamic equilibrium during the dissociation and association of the complex, the actual number may be even lower at any point of time. At the specified concentration of RecA protein in our experiment, relative to the concentration of the oligonucleotide, ATP hydrolysis is activated by less than 20% of the potentially maximum value (Figure 3a). Thus, the potential for RecA nucleation is severely limited. The addition of even a minute concentration of DprA dramatically elevates the rate of RecA-mediated ATP hydrolysis, regardless of the order of protein addition to the DNA. On the linear segment of RecA-dependent ATPase enhancement, it occurs proportionally to the amount of DprA protein added. In fact, saturation occurs at a DprA concentration of 0.3 μM, which is approximately one twentieth of the DNA concentration (5 μM). Taken together with the structural data, these results imply that for full nucleation stimulation, the DprA protein must interact with the DNA oligo along its entire length. In other words, it must interact or compete with every single monomer associated with the 34 nt oligos.

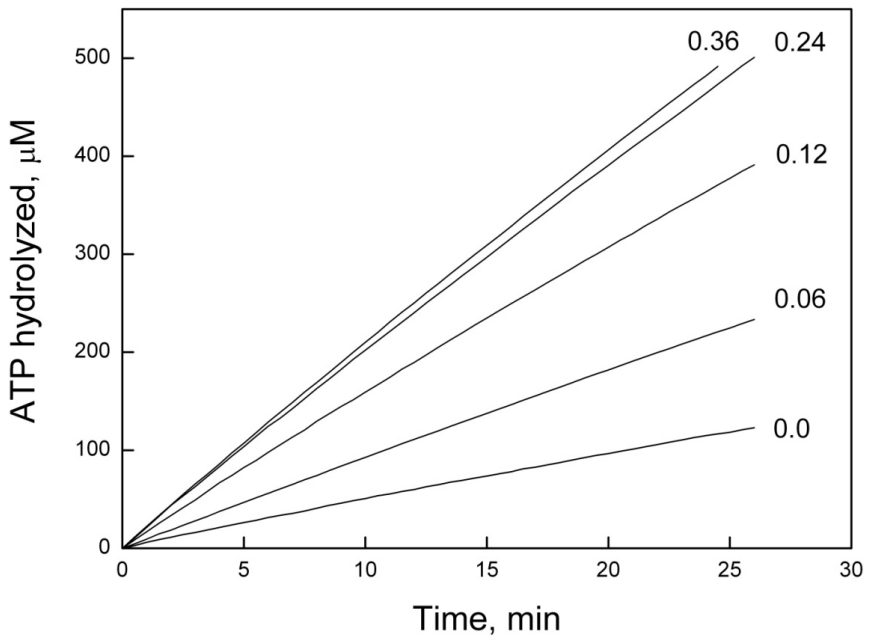

Fig. 3a In the presence of DprA protein, RecA protein can nucleate on the short DNA size. *The reactions were carried out as described under "Experimental Procedures" and contained 3 μm RecA protein, 5 μm oligonucletide, 2 mM ATP*

The DprA protein may interact with and facilitate the loading of RecA onto two types of substrates: those of suitable length for filament formation and those of shorter length. To more precisely determine the minimum substrate length allowing any stimulation of RecA nucleation by the DprA protein, we refined the experimental conditions. Adjustments were made to enhance RecA binding affinity for short oligonucleotides, removing filament length as a limiting factor in the RecA-DprA pair. This was achieved by substituting ATP for dATP in the hydrolysis assay, that is more consistent with the preferences of RecA protein from E.coli. Previous studies indicated that in the presence of dATP, filament disassembly slows down, and nucleation events on DNA are more frequent [25,26].

The use of dATP in this context is justified, as it reduces the number of monomers to a point where the complex length is solely limited by the DprA protein binding site size. Consequently, RecA-DprA nucleation on the oligo was examined in the presence of dATP. To assess the required length of ssDNA for stable interactions with RecA, dATPase assays were conducted using 50 nt, 34 nt and 21 nt oligos, with and without DprA protein (Figure 3b). The RecA protein failed to bind the 21 nt oligo alone but successfully bound the 50 nt oligo regardless of DprA protein presence. These data has demonstrated DprA's ability to specifically promote the loading of RecA onto substrates as short as 21 nt.

The 50-base length surpasses the critical threshold for sufficient RecA nucleation, rendering the contribution of DprA to the transition of RecA from a free state to a bound state visually as if undetectable. Although DprA remains present in the formation of long RecA filament, as indicated by the SSB protein displacement experiment, its stimulatory role becomes more evident with shorter oligonucleotides, increasing with oligo shortening and, consequently, filament length reduction.

The stimulatory effect of DprA appears to cease on lengths shorter than 18-21 bases, corresponding to the RecA hexamer's size. However, during the initial nucleation moment, the number of RecA monomers on DNA may be even smaller due to the DprA protein occupying its own DNA binding site, rendering this DNA inaccessible to other proteins. Short substrates may not provide enough binding sites or structural stability for DprA and RecA to interact effectively. These findings provide new insights into the established earlier model, which was based on the impact of oligonucleotide length on its interaction efficiency with DprA. In previous model, to assess this impact, the ssDNA binding properties of DprA protein were assayed using EMSA

techniques. Previous studies suggested that a homodimer DprA is required for optimal binding to ssDNA, setting the lower limit of oligonucleotide length around 40-50 bases [6, 15, 27].

Further experiments with a monomeric dimerization-deficient mutant confirmed that only the dimeric form is suitable for DNA binding [28]. Notably, these experiments excluded the presence of RecA protein during the subsequent step of its association with the complex. According to the presumed model, derived from structural analysis, a mechanism for RecA nucleation is extrapolated, despite the absence of experimental investigation through EMSA for this specific step. In line with this model, the subsequent step involves the interaction of the RecA monomer with the DprA-DprA-DNA complex, leading to the displacement of one of the DprA subunits or disruption of the dimeric structure [6]. Our dATPase experiments significantly complement this model by demonstrating that the DprA-RecA complex can form from the first stage, bypassing the "dimer subunit per two DNA sites" phase of interaction. Alternatively, if we assume that the dimer stage is indeed present, then the structure of the complex should be significantly revised, with the assumption of reducing the binding site size for DprA protein(s). However, with only biochemical experiments involving ATPase analysis and no structural studies, the picture cannot be made clearer. What can be said more definitively is that the DprA protein completely dissociates from the complex upon completion of RecA protein assembly on the oligonucleotide. Otherwise, under conditions of a lack of binding sites on DNA relative to the protein, the final rate of RecA-dependent ATPase should differ depending on the presence or absence of the DprA protein. Since the binding site of the DprA protein occupies almost half the length of the 21 nt oligonucleotide, such an influence on the ATPase rate would be difficult to ignore. It's challenging to definitively determine which model is more prevalent on long substrates, "dimer subunit per two DNA sites" or "dimer subunit per one DNA site". Current experiments cannot provide a conclusive answer.

Nevertheless, it is apparent that on a short substrate, such as 21 nucleotides, $DprA_2$-RecA complex binding becomes unfeasible. In this scenario, the simultaneous presence of the RecA protein should be sufficient to initially stabilize only one of the DprA monomers on DNA, and vice versa. Due to the noticeable overlap between the dimerization interfaces of DprA subunits and the interaction with RecA, both mechanisms represent possible scenarios. Nevertheless, we favor the hypothesis that the "dimer subunit per one DNA site" mechanism is more likely under certain favorable conditions, such as DNA length limitations by barriers like SSB protein or critical shifts in protein concentration balance.

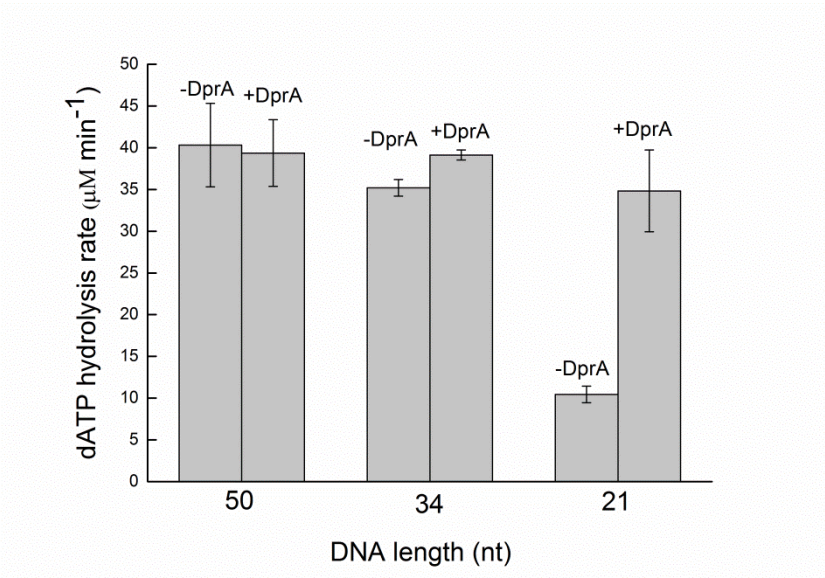

Fig. 3b Effect of DNA size on RecA-dependent ATP hydrolysis in the presence or absence of DprA protein. *The reactions were carried out as described under "Experimental Procedures" and contained 3 μm RecA protein, 5 μm oligonucletide, 2 mM dATP*

This fact inspired us to refine an existing model in which the structure of the complex could fit the experimental data.

Reconstruction of the structure of the RecA–DprA–ssDNA complex

The figure 4 shows a molecular model of the $RecA_3$–$(ATP–Mg^{2+})_2$–dT_{19} complex with the DprA protein dimer constructed on the basis of available structural data for three different complexes, which data on the spatial structure are published [28-30]. Although this nucleoprotein complex has a multicomponent structure, obtained from the structural data of complexes of various compositions, the structure and conformation of its constituent proteins turned out to be compatible with each other and does not have any structural problems. However, the conformation of the 5' end of the ssDNA presynaptic filament RecA (shown in brown) turned out to be incompatible with the presence of DprA proteins and directed towards the region of DprA protein-protein interface, and not towards the known nucleotide binding site found in the crystal structure of the dimeric complex of DprA proteins with short fragments of ssDNA [28]. This obviously means that binding of DprA proteins requires that the six nucleotide residues of the free 5' end of ssDNA adopt a suitable conformation for binding to the nucleotide binding site of the DprA protein. Molecular modeling of the conformational transition of ssDNA showed that the nucleotide from the dimeric complex of the DprA protein [28] and ssDNA of the presynaptic RecA protein filament [30] can be connected without any steric strains by two mobile nucleotide

residues, thus forming a continuous ssDNA chain running from the nucleotide binding site of the DprA protein to the RecA proteins helical filament as shown in the figure 4.

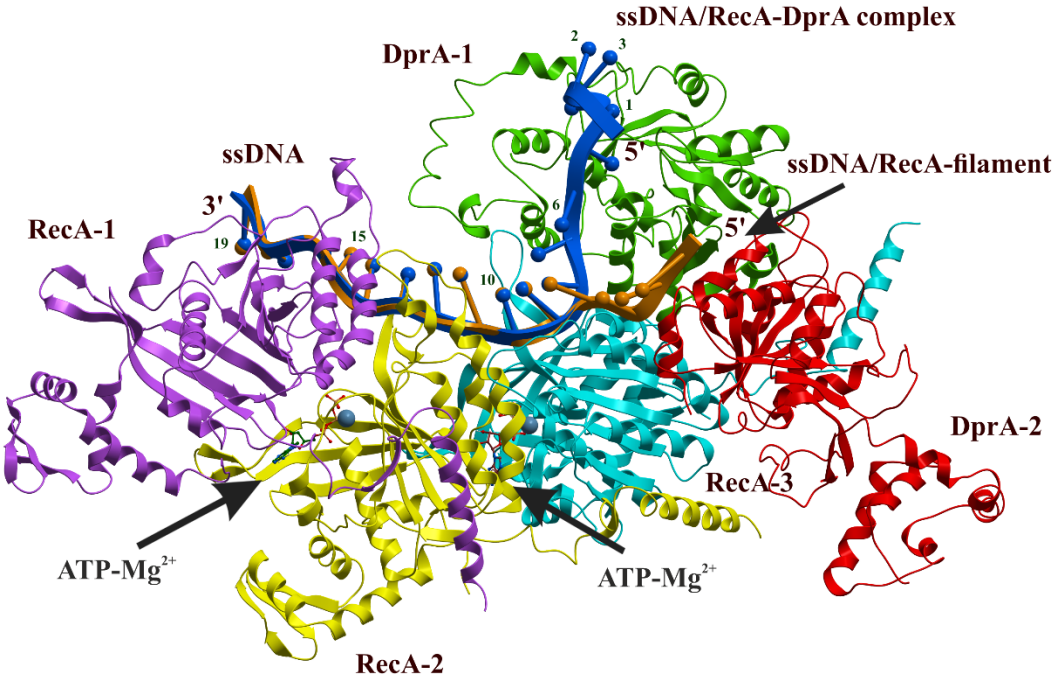

Figure 4. The molecular model of the complex of the RecA nucleoprotein filament ($RecA_3$–$(ATP–Mg^{2+})_2$–dT_{19})) and the DprA proteins. Three subunits of the RecA protein from *E.coli* are shown in blue, yellow and magenta respectively. Two DprA proteins from *E.coli* are shown in green and red. The conformation of the ssDNA (dT_{19}) where the conformation of the first six 5' terminal nucleotide residues is taken from the crystal structure of the ssDNA-DprA complex (PDB ID: 4LJR) [2] while the conformation of the last eleven 3' terminal nucleotide residues is taken from the crystal structure of the presynaptic RecA nucleoprotein filament (PDB ID: 3CMX) [1] is shown in blue. The original conformation of the ssDNA from the presynaptic RecA nucleoprotein filament is shown in brown for clarity. Two $ATP–Mg^{2+}$ substrate complexes in the RecA protein ATPase sites are shown in ball and stick representation. (PDB see in Supplemental Data)

The ATP hydrolysis site is located at the interface between two subunits of the RecA proteins. While the binding of RecA proteins to ssDNA in the filament structure requires the formation of direct contacts between loops L1 and L2 of neighboring RecA subunits every three nucleotide residues of ssDNA. Therefore, the minimum length of ssDNA required for all standard intermolecular interactions to be formed for at least one ATP binding site is 19 nucleotide residues as shown in the figure 4.

One of the interesting questions is the role of the second subunit of the DprA protein in the complex with $RecA_3$–$(ATP–Mg^{2+})_2$–$DprA_2$–dT_{19} shown in Figure 4. It is obvious that exactly the same symmetrical structure can be built on the basis of the complex of the second subunit of DprA with RecA protein. However, the simultaneous existence of these two structures in one complex is sterically impossible due to the competition of two symmetrically located subunits of the RecA protein for the same place in space.

Based on recently obtained structural data and the results of molecular modeling of complexes of the DprA protein with DNA and the RecA protein, a molecular mechanism of DprA-mediated loading of RecA proteins onto ssDNA has been proposed [31]. This molecular mechanism involves the formation of a dimeric nucleoprotein complex of DprA with ssDNA, which provides a suitable platform for further nucleation of the RecA filament, after which one of the DprA subunits leaves the complex and polymerization of RecA on ssDNA continues. It is assumed in the model that ssDNA binds simultaneously to two DprA subunits, the DNA binding sites of which are located at a distance of approximately 50Å, which corresponds to the length of ssDNA having 10 nucleotide residues in a fully unfolded extended conformation. It is known that the active center of DprA contains ssDNA with a length of 6 nucleotide residues [28]. In addition, the attachment of each new subunit of the RecA protein to the nucleoprotein filament requires three nucleotide residues of ssDNA [30]. Thus, the minimum requirements for the length of ssDNA for the formation of a complex with RecA capable of ATP hydrolysis is approximately equal to 6 + 10 + 6 + 2 + 11 = 35 nucleotide residues, which is significantly less than the estimate of 50 nucleotide residues obtained previously experimentally [6, 16].
It is also important that due to the symmetry of the dimeric complex of the DprA protein, the ssDNA chains in the ssDNA binding sites of the DprA proteins were directed in opposite directions. This makes it impossible to form a dimeric complex of DprA with one strand of ssDNA occupying both sites of the DprA proteins. Therefore, it is sterically possible to form of only a complex of DprA with ssDNA, occupying only one DprA protein ssDNA binding site as shown in Figure 4. And in this case, the minimum length of ssDNA is much shorter and amounts to 6+2+11=19 nucleotide residues. This estimate of the minimum required length of ssDNA explains the experimental data obtained in this work.

Discussion

In numerous bacterial species exhibiting natural transformation, dprA has been identified as a competence gene. The DprA protein has been implicated in safeguarding incoming DNA. Nonetheless, members of the dprA gene family (also known as smf) are detectable across virtually all bacterial species, implying broader functionality of their gene products. DprA proteins may undertake diverse roles in different bacterial contexts [2, 3]. We investigated the functionality of the DprA/Smf homolog in Escherichia coli in the context of Hfr conjugation. The overall recombination potential of bacterial cells is influenced by multiple factors. Recombination potential can be modulated by mutated RecA or various other recombination functions that support RecA, such as MutS, RecX, DinI, or RecFOR [13,15, 32,33]. While most regulators only partially alter the genetic outcomes of recombination, thereby adjusting the

natural FRE values used to assess in vivo recombinase activity, our study reveals a striking observation under normal conditions: DprA emerges as a potent activator of recombination activity in vivo, significantly augmenting FRE upon overexpression (by approximately 18-fold). Thus far, we have not succeeded in demonstrating a similar effect for any of the other regulatory proteins. When expressed in E. coli, DprA B.s complements a dprA null mutant in E. coli for Hfr conjugation. These observations seem to indicate that the function of DprA extends beyond facilitating natural transformation and may serves a role in the fitness of bacterial species that are not typically competent for DNA uptake. Interestingly, homologs of DprA are also identified in species like Mycobacterium tuberculosis, among others. This species is believed to have evolved strictly through clonal descent and exhibits no evidence of horizontal gene transfer [34]. E. coli, as known, harbors two major recombination systems termed RecBCD and RecFOR, which appear to operate during conjugational recombination [35,36]. It is conceivable, that the DprA–RecA pathway in E. coli is likely to be regarded as the third recombination pathway, the exploration of which will contribute to a more comprehensive understanding of homologous recombination.

During conjugation, it is conventionally assumed that one significant barrier to RecA protein binding to ssDNA is the presence of the SSB protein. The RecA-DprA complex exhibited an enhanced capacity to displace the EcSSB protein. This study, however, was carried out using heterologous proteins (DprAB.c, RecAE.c, and SSBE.c). Our ATPase analysis, where DprA and RecA are simultaneously involved, allowed us to perform several clarifying experiments, the main one being the search for the minimum length of substrate capable of providing a platform for nucleation of the RecA filament. The conclusion is that the real length of DNA required for a complex containing both proteins is only about twenty nucleotides. This fact inspired us to refine an existing model in which the structure of the complex could fit the experimental data. The structural and biochemical characterisation of the DprA protein reported here showed that it is sterically possible to form only a complex of DprA with ssDNA where DprA occupies only one binding site with ssDNA. Given the known size of the DprA-DNA binding site, we extrapolate these data to construct a ternary RecA -DprA -ssDNA complex that occurs during the RecA filament nucleation. Within this context, the remaining olig space can accommodate, at most, a RecA trimer as the nucleation unit, which subsequently fills the rest of the substrate upon DprA dissociation, forming a hexamer. The complex structure between DprA and ssDNA reported here provides us with a good starting point for further studies on DprA and other related proteins. Undoubtedly, regulators of homologous recombination involved in the processes of chromosomal transformation should be considered as potential targets for inhibition [37, 38].

Further study of the structure of the DprA-RecA complex may lead to the design of compounds or inhibitors of horizontal gene transfer.

Material and Methods

Strains and plasmids

Donor KL227 (HfrP4x metB) and recipients: AB1157 (thr-1 leuB6 ara14 proA2 hisG4 argE3 thi-1 supE44 rpsL31) and recombination deficient JC10289 (as AB1157 but Δ[recA-srlR306]::Tn10 = ΔrecA306) were from A.J. Clark's collection. Strain JW5708-1 (Δsmf-724::kan) was from Keio Collection (E.coli Genetic Resources at Yale CGSC). The dprA-deficient strains were constructed by P1 transduction to transfer smf-724 (del)::kan from strain JW5708-1 into AB1157 as described previously [39]. Plasmids pdprABc was constructed and kindly supplied by Prof. J.C. Alonso. Plasmid pdprAEc was constructed and supplied by Eurogen. These plasmids contain the pdprAEc and pdprABc genes under a T7 RNA polymerase driven promoter.

Proteins

The wild-type E. coli RecA was purified as previously described [40]. Single-strand binding (SSB) protein was kindly provided by Prof. M. Cox (University of Wisconsin-Madison) Their concentrations were determined using native extinction coefficients: $\epsilon 280 = 2.23 \times 10^4\ M^{-1}cm^{-1}$ for RecA protein [41], and $\epsilon 280 = 2.38 \times 10^4\ M^{-1}cm^{-1}$ for SSB protein [42]. The concentration of B. subtilis DprA was determined from the absorbance at 280 nm using the native extinction coefficient $4.5 \times 10^4\ M^{-1}cm^{-1}$ [7]

B. subtilis DprA was purified similarly to the previously described protocol with minor modifications [7]. Briefly, BL21(DE3)pLysE cells were transformed with the pCB888 plasmid kindly provided by prof. J.C. Alonso. Bacterial cells were grown at 25°C and expression was induced by the addition of 0.4 mM isopropyl-1-thio-β-D-galactopyranoside (IPTG) at an OD600 of 0.5. After induction, cells were grown for 3 hours and rifampicin (200 µg/ml) was added 30 minutes after IPTG addition. DprA was purified from the clarified lysate by two-step chromatography using HisTrap HP 1 mL (GE Healthcare) and HiTrap SP XL 1 mL columns (GE Healthcare). Fractions containing DprA were concentrated on an Amicon Ultra-4 centrifugal filter (Merck) with a 3 kDa cut-off, supplemented with 50% glycerol and stored at -20°C.

ATP hydrolysis (ATPase) Assays

A coupled enzyme, spectrophotomeric assay [43,44] was used to measure RecA-mediated ATP hydrolysis. The ADP generated by hydrolysis was converted back to ATP by a regeneration system of pyruvate kinase and phosphoenolpyruvate (PEP). The resultant pyruvate was

converted to lactate by lactate dehydrogenase using NADH as a reducing agent. The conversion of NADH to NAD+ was monitored as a decrease in absorbance at 380 nm. The amount of ATP hydrolyzed over time was calculated using the NADH extinction coefficient ε380 = 1.21 $mM^{-1}cm^{-1}$. The assays were carried out in a Cary 5000 dual beam spectrometer, with a temperature controller and a 12-position cell changer. The path length was 1 cm, the band pass was 2 nm. All ATPase assays contained a reaction solution of 25 mM Tris.HCl (pH 7.5, 88% cation), 10 mM MgCl, 5% w/v glycerol, 1 mM dithiothreitol (DTT), 3 mM PEP, 10 U/ml pyruvate kinase, 10 U/ml lactate dehydrogenase, 4.5 mM NADH and 5 μM M13mp18 cssDNA. All reactions were repeated at least three times with similar results.

Conjugation

Conjugation was carried out essentially as described [13]. Both Hfr and F−strains were grown, crossed and selected for recombinants at 37°C in mineral salts 56/2 medium supplied with all necessary growth factors at pH 7.5. The ratio between donors and recipients in the mating mixture was 1:10, 2–4 x107 donors and 2–4 x 108 recipients per 1 ml. The yield of Thr+Strr and Ara+Strr recombinants in all independent crosses (5–7% relative to donors) was normalized according to the mating ability of each recipient used. The latter was determined by the yield of transconjugants F'-lac+ in crosses between the recipients and donor P200 F'-lac.

FRE value calculations were carried out as described [13]. Alterations in FRE (ΔFRE) promoted by the normal and truncated PaRecA gene or by the EcRecA gene with appropriate modifications relative to the FRE value promoted by the wild type EcRecA gene were calculated using the following formula: ΔFRE = ln(2μ1 - 1)/(2 μ2 - 1), where μ1 is the linkage of selected thr+ or ara+ and unselected leu+ markers in a cross using wild type E. coli strain AB1157 and μ2 is the similar linkage in the cross being analyzed. Calculations of uncertainty of relative FRE values were determined as deviations from the average values by making use of the program Excel-97 with formula [=2*STDEV] and by inputting the values from independent repeats of three experiments.

Reconstruction of the structure of the RecA–DprA–ssDNA complex

To construct a full-atomic model of the spatial structure of the RecA–DprA–ssDNA complex, we used a set of crystal structures from the Protein Data Bank, namely the structure of the presynaptic nucleoprotein filament of the RecA protein containing $RecA_6$–($ADP–AlF_4–Mg)_6$–dT_{18} (PDB ID: 3CMX) [30], the structure of the DprA protein complex from Helicobacter pylori with ssDNA ($DprA–dT_6)_2$ [2] (PDB ID: 4LJR) and the spatial model of the RecA-DprA complex obtained using molecular docking methods, which was kindly provided by the authors of this work [29]. The structure superimposition, energy minimization and other molecular

manipulations were carried out using standard protocols of the ICM-Pro software package (Molsoft LLC, USA), [45] and the ECEPP/3 force field as implemented in the ICM-Pro [45, 47, 48].

The spatial structure of the dimeric complex of the DprA proteins from E.coli was obtained based on the crystal structure of the DprA dimeric complex from Helicobacter pylori (PDB ID: 4LJR) using standard homology modeling protocols implemented in ICM-Pro. The RecA–DprA–ssDNA complex was assembled using a series of superimpositions of the spatial structure of proteins. At the first stage, the position of the DprA dimer was determined using the superimposition of RecA monomers in the presynaptic nucleoprotein filament (PDB ID: 3CMX) and in the spatial model of the RecA-DprA complex [29]. The position of 6 nucleotides at the 5' end of ssDNA was then determined using superposition of the DprA proteins in the complex (DprA-dT_6)$_2$ (PDB ID: 4LJR) and the molecular model of the RecA-DprA complex [29], while the conformation of the 3'-terminal part of ssDNA was taken in the form in which it is located in the presynaptic nucleoprotein filament RecA (PDB ID: 3CMX). Finally, the conformation of 2 bridging nucleotides was determined by minimizing the energy of ssDNA with fixed 3' and 5' ssDNA ends.

Funding

The research was supported by Russian Science Foundation, project No. 22-74-00072 to NM. The research was carried out within the state assignment of Ministry of Science and Higher Education of the Russian Federation (theme №1023031500033-1-1.6.7;1.6.4;1.6.8) to IB and MP. The work was carried out using scientific equipment of the Center of Shared Usage "The analytical center of nano- and biotechnologies of SPbPU".

Acknowledgements

We are grateful to Prof. Juan C. Alonso (Department of Microbial Biotechnology, Centro Nacional de Biotecnología, Madrid, Spain) for his generous gift of expression plasmids bearing the B. subtilis DprA for the E. coli strain. The authors would like to thank Prof. Michael M. Cox (University of Wisconsin–Madison) for SSB protein. The authors thank Jessica Andreani and Sophie Quevillon-Cheruel (Institute of Integrative Biology of the Cell, Université Paris-Sud, Gif-sur-Yvette, France) for structural model of the complex of the DprA dimer with RecA protein. We acknowledge the help of Inna Kurdyumova, Daria Antonova and Maryam Dmitrieva (Peter the Great St. Petersburg Polytechnic University) in purification of B. subtilis DprA.

References

1. Johnston C, Mortier-Barriere I, Khemici V, Polard P (2018) Fine-tuning cellular levels of DprA ensures transformant fitness in the human pathogen Streptococcus pneumoniae. *Mol Microbiol* 109, 663-675.
2. Johnston C, Martin B, Fichant G, Polard P, Claverys JP (2014) Bacterial transformation: distribution, shared mechanisms and divergent control. *Nat Rev Microbiol* 12, 181-96.
3. Brito IL (2021) Examining horizontal gene transfer in microbial communities. *Nat Rev Microbiol* 19, 442-453.

4. Takata T, Ando T, Israel DA, Wassenaar TM, Blaser MJ (2005) Role of dprA in transformation of Campylobacter jejuni. *FEMS Microbiol Lett* 252, 161-8.
5. Mortier-Barrière I, Velten M, Dupaigne P, Mirouze N, Piétrement O, McGovern S, Fichant G, Martin B, Noirot P, Le Cam E, Polard P, Claverys JP (2007) A key presynaptic role in transformation for a widespread bacterial protein: DprA conveys incoming ssDNA to RecA. *Cell* 130, 824-36.
6. Quevillon-Cheruel S, Campo N, Mirouze N, Mortier-Barrière I, Brooks MA, Boudes M, Durand D, Soulet AL, Lisboa J, Noirot P, Martin B, van Tilbeurgh H, Noirot-Gros MF, Claverys JP, Polard P (2012) Structure-function analysis of pneumococcal DprA protein reveals that dimerization is crucial for loading RecA recombinase onto DNA during transformation. *Proc Natl Acad Sci U S A* 109, 2466-75.
7. Yadav T, Carrasco B, Hejna J, Suzuki Y, Takeyasu K, Alonso JC (2013) Bacillus subtilis DprA recruits RecA onto single-stranded DNA and mediates annealing of complementary strands coated by SsbB and SsbA. *J Biol Chem* 288, 22437-50.
8. Ithurbide S, Coste G, Lisboa J, Eugénie N, Bentchikou E, Bouthier de la Tour C, Liger D, Confalonieri F, Sommer S, Quevillon-Cheruel S, Servant P (2020) Natural Transformation in *Deinococcus radiodurans*: A Genetic Analysis Reveals the Major Roles of DprA, DdrB, RecA, RecF, and RecO Proteins. *Front Microbiol* 11, 1253.
9. Sharma DK, Misra HS, Soni I, Rajpurohit YS (2022) Characterization of DNA Processing Protein A (DprA) of the Radiation-Resistant Bacterium Deinococcus radiodurans. *Microbiol Spectr* 10, e0347022.
10. Carrasco B, Manfredi C, Ayora S, Alonso JC (2008) Bacillus subtilis SsbA and dATP regulate RecA nucleation onto single-stranded DNA. *DNA Repair (Amst)* 7, 990-6.
11. Smeets LC, Becker SC, Barcak GJ, Vandenbroucke-Grauls CM, Bitter W, Goosen N (2006) Functional characterization of the competence protein DprA/Smf in Escherichia coli. *FEMS Microbiol Lett* 263, 223-8.
12. Lawley TD, Gordon GS, Wright A, Taylor DE (2002) Bacterial conjugative transfer: visualization of successful mating pairs and plasmid establishment in live Escherichia coli. *Mol Microbiol* 44, 947-56.
13. Lanzov VA, Bakhlanova IV, Clark AJ (2003)Conjugational hyperrecombination achieved by derepressing the LexA regulon, altering the properties of RecA protein and inactivating mismatch repair in Escherichia coli K-12. *Genetics* 163, 1243-54.
14. Lloyd RG (1978) Hyper-recombination in Escherichia coli K-12 mutants constitutive for protein X synthesis. *J Bacteriol* 134, 929-35.
15. Bakhlanova IV, Dudkina AV, Wood EA, Lanzov VA, Cox MM, Baitin DM (2016) DNA Metabolism in Balance: Rapid Loss of a RecA-Based Hyperrec Phenotype. *PLoS One*. 11, e0154137.
16. Beyene GT, Kalayou S, Riaz T, Tonjum T (2017) Comparative proteomic analysis of Neisseria meningitidis wildtype and dprA null mutant strains links DNA processing to pilus biogenesis. *BMC Microbiol* 17, 96.
17. Yadav T, Carrasco B, Serrano E, Alonso JC (2014) Roles of Bacillus subtilis DprA and SsbA in RecA-mediated genetic recombination. *J Biol Chem* 289, 27640-52.
18. Arenson TA, Tsodikov OV, Cox MM (1999) Quantitative analysis of the kinetics of end-dependent disassembly of RecA filaments from ssDNA. *J Mol Biol* 288, 391-401.
19. Bork JM, Cox MM, Inman RB (2001) RecA protein filaments disassemble in the 5' to 3' direction on single-stranded DNA. *J Biol Chem* 276, 45740-3.
20. Hovland E, Beyene GT, Frye SA, Homberset H, Balasingham SV, Gómez-Muñoz M, Derrick JP, Tønjum T, Ambur OH. 2017 DprA from Neisseria meningitidis: properties and role in natural competence for transformation. *Microbiology (Reading)* 163, 1016-1029.
21. Brenner SL, Mitchell RS, Morrical SW, Neuendorf SK, Schutte BC, Cox MM (1987) recA protein-promoted ATP hydrolysis occurs throughout recA nucleoprotein filaments. *J Biol Chem* 262, 4011-6.

22. Sussman R, Sharma SK, Kuzirian A (2008) Catalytic activities of recA protein are dependent on the lattice length of the single-strand DNA ligand. *Cell Cycle* 7, 89-95.
23. Galletto R, Amitani I, Baskin RJ, Kowalczykowski SC (2006) Direct observation of individual RecA filaments assembling on single DNA molecules. *Nature* 443, 875-8.
24. Bell JC, Plank JL, Dombrowski CC, Kowalczykowski SC (2012) Direct imaging of RecA nucleation and growth on single molecules of SSB-coated ssDNA. *Nature* 491, 274-8.
25. Menetski JP, Kowalczykowski SC (1989) Enhancement of Escherichia coli RecA protein enzymatic function by dATP. *Biochemistry* 28, 5871-81.
26. Shan Q, Cox MM (1996) RecA protein dynamics in the interior of RecA nucleoprotein filaments. *J Mol Biol* 257, 756-74.
27. Dwivedi GR, Sharma E, Rao DN (2013) Helicobacter pylori DprA alleviates restriction barrier for incoming DNA. *Nucleic Acids Res* 41, 3274-88.
28. Wang W, Ding J, Zhang Y, Hu Y, Wang DC (2014) Structural insights into the unique single-stranded DNA-binding mode of Helicobacter pylori DprA. *Nucleic Acids Res* 42, 3478-91.
29. Lisboa, J., Andreani, J., Sanchez, D., Boudes, M., Collinet, B., Liger, D., van Tilbeurgh, H., Guerois, R. and Quevillon-Cheruel, S. (2014) Molecular determinants of the DprA-RecA interaction for nucleation on ssDNA. *Nucleic Acids Res*, **42**, 7395-7408.
30. Chen, Z., Yang, H. and Pavletich, N.P. (2008) Mechanism of homologous recombination from the RecA-ssDNA/dsDNA structures. *Nature*, **453**, 489-484.
31. Sharma, D.K., Misra, H.S., Bihani, S.C. and Rajpurohit, Y.S. (2023) Biochemical Properties and Roles of DprA Protein in Bacterial Natural Transformation, Virulence, and Pilin Variation. *Journal of Bacteriology*, **205**, e00465-00422.
32. Baitin DM, Bakhlanova IV, Chervyakova DV, Kil YV, Lanzov VA, Cox MM (2008) Two RecA protein types that mediate different modes of hyperrecombination. *J Bacteriol* 190, 3036-45.
33. Bakhlanova IV, Dudkina AV, Baitin DM, Knight KL, Cox MM, Lanzov VA (2010) Modulating cellular recombination potential through alterations in RecA structure and regulation. *Mol Microbiol* 78, 1523-38.
34. Supply P, Warren RM, Bañuls AL, Lesjean S, Van Der Spuy GD, Lewis LA, Tibayrenc M, Van Helden PD, Locht C (2003) Linkage disequilibrium between minisatellite loci supports clonal evolution of Mycobacterium tuberculosis in a high tuberculosis incidence area. *Mol Microbiol* 47, 529-38.
35. Clark AJ, Sandler SJ (1994) Homologous genetic recombination: the pieces begin to fall into place. *Crit Rev Microbiol* 20, 125-42.
36. Amundsen SK, Smith GR (2003) Interchangeable parts of the Escherichia coli recombination machinery. *Cell* 112, 741-4.
37. Yakimov A, Pobegalov G, Bakhlanova I, Khodorkovskii M, Petukhov M, Baitin D (2017) Blocking the RecA activity and SOS-response in bacteria with a short α-helical peptide. *Nucleic Acids Res* 45, 9788-9796.
38. Yakimov A, Bakhlanova I, Baitin D (2021)Targeting evolution of antibiotic resistance by SOS response inhibition. *Comput Struct Biotechnol J* 19, 777-783.
39. Csonka LN, Clark AJ (1980) Construction of an Hfr strain useful for transferring recA mutations between Escherichia coli strains. *J Bacteriol* 143, 529-30.
40. Cox MM, McEntee K, Lehman IR (1981) A simple and rapid procedure for the large scale purification of the recA protein of Escherichia coli. *J Biol Chem* 256, 4676-8.
41. Drees JC, Lusetti SL, Chitteni-Pattu S, Inman RB, Cox MM (2004) A RecA filament capping mechanism for RecX protein. *Mol Cell* 15, 789-98.
42. Lohman TM, Green JM, Beyer RS (1986) Large-scale overproduction and rapid purification of the Escherichia coli ssb gene product. Expression of the ssb gene under lambda PL control. *Biochemistry* 25, 21-5.

43. Morrical SW, Lee J, Cox MM (1986) Continuous association of Escherichia coli single-stranded DNA binding protein with stable complexes of recA protein and single-stranded DNA. *Biochemistry* 25, 1482-94.
44. Lindsley JE, Cox MM (1990) Assembly and disassembly of RecA protein filaments occur at opposite filament ends. Relationship to DNA strand exchange. *J Biol Chem* 265, 9043-54.
45. Abagyan, R. and Totrov, M. (1994) Biased probability Monte Carlo conformational searches and electrostatic calculations for peptides and proteins. *J Mol Biol*, **235**, 983-1002.
46. Arnautova, Y.A., Jagielska, A. and Scheraga, H.A. (2006) A new force field (ECEPP-05) for peptides, proteins, and organic molecules. *J Phys Chem B*, **110**, 5025-5044.
47. Nemethy, G., Gibson, K.D., Palmer, K.A., Yoon, C.N., Paterlini, G., Zagari, A., Rumsey, S. and Scheraga, H.A. (1992) Energy Parameters in Polypeptides .10. Improved Geometrical Parameters and Nonbonded Interactions for Use in the Ecepp/3 Algorithm, with Application to Proline-Containing Peptides. *J. Phys. Chem.*, **96**, 6472-6484.
48. Nemethy, G., Pottle, M.S. and Scheraga, H.A. (1983) Energy Parameters in Polypeptides .9. Updating of Geometrical Parameters, Nonbonded Interactions, and Hydrogen-Bond Interactions for the Naturally-Occurring Amino-Acids. *J. Phys. Chem.*, **87**, 1883-1887.