The Lon protease-like domain in the bacterial RecA paralog RadA is required for DNA binding and repair

Received for publication, November 29, 2016, and in revised form, April 16, 2017 Published, Papers in Press, April 21, 2017 DOI 10.1074/jbc.M116.770180

Masao Inoue‡, Kenji Fukui‡, Yuki Fujii‡, Noriko Nakagawa‡, Takato Yano§, Seiki Kuramitsu¶, and Ryoji Masui‡

From the ‡Department of Biological Sciences, Graduate School of Science, Osaka University, 1-1 Machikaneyama-cho, Toyonaka, Osaka 560-0043, the §Department of Biochemistry, Osaka Medical College, 2-7 Daigakumachi, Takatsuki, Osaka 569-8686, the ¶Graduate School of Frontier Biosciences, Osaka University, 1-3 Yamadaoka, Suita, Osaka 565-0871, and the ‡Graduate School of Science, Osaka City University, 3-3-138 Sugimoto, Sumiyoshi-ku, Osaka 558-8585, Japan

Edited by Patrick Sung

Homologous recombination (HR) plays an essential role in the maintenance of genome integrity. RecA/Rad51 paralogs have been recognized as an important factor of HR. Among them, only one bacterial RecA/Rad51 paralog, RadA, is involved in HR as an accessory factor of RecA recombinase. RadA has a unique Lon protease-like domain (LonC) at its C terminus, in addition to a RecA-like ATPase domain. Unlike Lon protease, RadA’s LonC domain does not show protease activity but is still essential for RadA-mediated DNA repair. Reconciling these two facts has been difficult because RadA’s tertiary structure and molecular function are unknown. Here, we describe the hexameric ring structure of RadA’s LonC domain, as determined by X-ray crystallography. The structure revealed the two positively charged regions unique to the LonC domain of RadA are located at the intersubunit cleft and the central hole of a hexameric ring. Surprisingly, a functional domain analysis demonstrated the LonC domain of RadA binds DNA, with site-directed mutagenesis showing that the two positively charged regions are critical for this DNA-binding activity. Interestingly, only the intersubunit cleft was required for the DNA-dependent stimulation of ATPase activity of RadA, and at least the central hole was essential for DNA repair function. Our data provide the structural and functional features of the LonC domain and their function in RadA-mediated DNA repair.

Homologous recombination (HR) is one of the most important biological phenomena for maintaining genome integrity through DNA-processing pathways, such as double-strand break repair and the rescue of stalled replication forks (1–4). HR also generates genetic variation and is recognized as one of the driving forces of evolution through horizontal gene transfer or intrachromosomal recombination (5, 6). To fulfill these contradictory roles, numerous proteins engage in each reaction in various organisms, such as *Escherichia coli*, *Bacillus subtilis*, and *Deinococcus radiodurans* (14–22). Recently, a biochemical characterization of *E. coli* RadA protein has shown that RadA stimulates RecA-catalyzed strand exchange *in vitro* (23). RadA is a multidomain protein consisting of three putative functional regions (Fig. 1A). A zinc-finger (ZF) motif, a RecA-like ATPase domain, and a Lon protease-like domain (LonC) are located at the N-terminal, central, and C-terminal regions, respectively. The ZF and LonC domains are unique to the bacterial RecA paralogs and are not found in eukaryotic and archaeal RecA/Rad51 paralogs.

A LonC domain was originally identified in Lon protease as a catalytic domain of protease activity (24–26). Lon protease is a member of the ATP-dependent serine proteases found in three kingdoms of life and involved in protein degradation upon various stress responses (27, 28). It has been thought that a protein substrate binds to an N-terminal putative substrate recognition domain, is unfolded by a central AAA⁺ motor domain, and then is degraded by the C-terminal LonC domain (Fig. 1A). Structural and mutational analyses of the LonC domain from *E. coli* Lon protease have shown that the protease activity of the LonC domain is exerted by a catalytic Ser residue (24, 25). Interestingly, different from Lon proteases, the catalytic Ser is not conserved in a LonC domain of RadA, and it has been shown that RadA has no protease activity (14). As a matter of convenience, we refer to a LonC domain of the Lon protease and that of RadA as pLonC (Lon protease-type) and rLonC (RadA-type), respectively. Genetic studies have indicated that rLonC has an essential role in RadA-mediated DNA repair (16, 22). However, the tertiary structure and molecular function of rLonC are still unknown.

In this study, we determined the X-ray crystal structure of rLonC from *Thermus thermophilus* HB8 and revealed that rLonC possesses the DNA-binding activity exerted by two pos-
Novel Lon protease-like DNA-binding domain of RecA paralog

E. coli Lon protease, Ser-679 is a catalytic residue of the central hole of the hexameric ring structure. We also showed that these two DNA-binding sites have different roles in the DNA-dependent ATPase activity of RadA in vitro and that at least the central hole is essential to the RadA-mediated DNA repair function in vivo. Furthermore, combining these results with a phylogenetic analysis highlighted structural and functional relationships between RadA and another LonC family protein, ComM. Our findings provide not only the structural and functional aspects of the rLonC domain in RadA-mediated DNA repair but also evolutionary aspects of the LonC family proteins.

Results
Overall structure of rLonC
First, we performed structural prediction and multiple sequence alignment analyses to determine the domain organization of RadA homologs (Fig. 1A). Structural prediction of T. thermophilus HB8 RadA (423 residues) using the SWISS-MODEL (29) and the I-TASSER (30) servers showed that a region of amino acid residues 256–275 of RadA from T. thermophilus HB8 could be a linker region between the RecA-like ATPase domain and rLonC. Based on this prediction, we prepared rLonC (amino acid residues 262–423) in N-terminal His6-tagged form for X-ray crystallographic analysis. The crystal structure of rLonC was determined at 2.7 Å resolution by molecular replacement using the structural model predicted by the I-TASSER server (30) as a search model. The statistics for the data collection and model refinement are summarized in Table 1.

The overall structure of rLonC was composed of a core structure, ribosomal protein S5 domain 2-type βββαβαβα fold, and an extra C-terminal α/β subdomain (Fig. 1B). The structure of rLonC determined in this study was similar to that of pLonC from E. coli (25) as denoted by a root mean square (r.m.s.) deviation of 3.9 Å for Cα atoms. As shown in Fig. 1, B and C, pLonC has two α-helices and one β-sheet in the extra C-terminal α/β subdomain (25), although these elements were absent in the rLonC structure. It should be noted that the N-terminal His6 tag and 10 amino acid residues exhibited high B-factor values or were disordered in both molecules in the asymmetric unit. The N-terminal disordered region of one molecule in the asymmetric unit was longer (residues 262–271) than that of the other molecule (residues 262–266) by five residues.

In E. coli Lon protease, Ser-679 is a catalytic residue of the protease activity and forms a catalytic dyad with Lys-722 (25). However, in the rLonC structure, Gly-350 and Arg-393 were located at the positions corresponding to Ser-679 and Lys-722 of E. coli Lon protease, respectively (Fig. 1, B and C). Furthermore, there was no serine residue around Gly-350 or Arg-393. These findings are consistent with a previous report that RadA had no protease activity (14).

Oligomeric structure of rLonC
In a crystal, rLonC forms a hexameric ring structure related by crystallographic 3-fold symmetry of a non-crystallographic dimer with an r.m.s. deviation of 0.16 Å for Ca atoms (Fig. 1D). The subunit arrangement of the hexameric ring structure of rLonC was similar to that of E. coli pLonC (25). In addition, like E. coli pLonC, the dominant interactions in the subunit interface are hydrophilic, including 11 hydrogen bonds and 6–7 salt bridges in each subunit interface. The total buried accessible surface area of rLonC upon hexamer formation was 10,464 Å², comparable with the 9,169 Å² for E. coli pLonC (25). The hexameric ring was predicted to be the most stable oligomer in the crystal structure by the PISA program (31). By contrast, gel-filtration analyses revealed that in solution the rLonC domain and full-length RadA existed as trimers to tetramers and dimers to pentamers, respectively (Fig. 2). In addition, the dissociation of the oligomeric structure was observed in a salt-dependent manner. These behaviors in the oligomer formation of RadA are similar to those of Lon proteases, in that they form various types of oligomers (32–35).

Unlike in E. coli pLonC (Fig. 1E, top right panel), the surface electrostatic potential map showed that rLonC had positively charged regions at one side of the ring (Fig. 1E, top left panel). At this “basic” side of the ring, a cleft-like structure is formed at the subunit interface (hereafter referred to as the “intersubunit cleft”) (Fig. 1E, bottom left panel). At the other side of the ring, there was no significant biased distribution of charged residues (Fig. 1E, bottom right panel). In addition, rLonC had another positively charged region at the central hole of the ring (hereafter referred to as the “central hole”) (Fig. 1E, top left panel). On the basic side, the inner diameter of the entrance of the central hole is ~29 Å. The hole is gradually narrower from the basic side to the other side of the ring, and the inner diameter at the exit formed by the loop containing 290TPFPA295 is ~17 Å (Fig. 1D, left panel). On the basis of these unique features, especially the positively charged regions, of rLonC’s structure, we hypothesized that rLonC has a DNA-binding ability.

rLonC is a novel DNA-binding module
To verify the hypothesis that rLonC has DNA-binding activity, we constructed and purified various domain-truncated mutants of RadA and performed an electromobility shift assay (EMSA) using 60 bp of double-stranded DNA (dsDNA) (Fig. 3, A and B). As expected, the mutants containing the rLonC domain exhibited DNA-binding activity. In the presence of the rLonC domain, as the protein concentration increased, bands of the DNA–protein complex appeared, and those of free DNA disappeared (Fig. 3A). Binding of the wild type, rLonC (∆N261) and ∆N53 to DNA, showed positive cooperativity, and the dissociation constants of the wild type and rLonC were comparable (Fig. 3B and Table 2). However, the loss of the N-terminal ZF domain (∆N53) resulted in considerably lower DNA-binding activity than the wild type.

Interestingly, our analysis could not detect the DNA-binding activity in either the N-terminal ZF domain or the RecA-like ATPase domain (Fig. 3, A and B). Structural prediction and multiple sequence alignment analyses support this finding in the RecA-like ATPase domain of RadA (Fig. 3, C and D). The L1 and L2 loops of RecA have been thought to be important for DNA binding (36, 37). In the case of RadA, the corresponding loops are expected to be shorter than those of RecA, and the residues that are important for DNA binding are not conserved (Fig. 3, C and D). However, there is still the possibility that the
DNA-binding activities of the ZF and the ATPase domains depend on the presence of the rLonC domain. Further analysis is required to elucidate whether the ZF domain and the ATPase domain have DNA-binding ability.

It has been shown that Lon protease activity is modulated by DNA binding and that the DNA-binding ability resides in the \( \text{H9251} \)-subdomain of the N-terminal AAA \( \text{H11001} \)-domain but not in the pLonC domain (32, 38, 39). As mentioned above, rLonC has unique positively charged regions that are not found in pLonC. Thus, we expected that the positively charged regions at the surface of rLonC's structure have an important role in the DNA-binding ability.

Identification of the DNA-binding residues

To identify the residues responsible for DNA binding, we selected 11 basic surface residues (Fig. 4A) and performed ala-
Novel Lon protease-like DNA-binding domain of RecA paralog

Table 1
Data collection and refinement statistics

| rLonC       | rLonC       |
|-------------|-------------|
| Data collection* | Data collection* |
| Beamline     | SPring-8 BL38B1 |
| Space group  | P4,32 |
| Cell dimensions | 156.29, 156.29, 156.29 |
| α, β, γ (°)  | 90.00, 90.00, 90.00 |
| Wavelength (Å) | 1.00000 |
| Resolution (Å) | 49.4–2.70 (2.75–2.70) |
| I/α(I)       | 0.051 (0.486) |
| Rfree,Rwork  | 20.4 (1.9) |
| CC1/2        | 0.625 |
| Completeness (%) | 100 (100) |
| Redundancy   | 41.3 (41.6) |

* One crystal was used for each data set.

The values in parentheses are for the highest resolution shell.

a Resolution is defined as 2π/λ, where λ is the wavelength of the incident x-rays.

b Rfree = \sum hkl (|Fo| - |Fcalc|)^2/\sum hkl |Fo|^2, where \sum hkl is the intensity of all reflections with index hkl.

c Rwork = \sum hkl (|Fo| - |Fcalc|)^2/\sum hkl |Fo|^2, where \sum hkl is the intensity of all reflections with index hkl.

d The values of the highest resolution shell are shown.

Figure 2. Gel filtration analysis of the rLonC domain and full-length RadA. The elution profiles of the rLonC domain (ΔN261) and full-length RadA (WT) in various salt concentrations are shown in the top and bottom panels, respectively. The NaCl concentrations are labeled. The NaCl concentrations are labeled. The elution peaks and molecular masses of standard proteins are indicated as black arrows and labeled in kDa, respectively. The estimated elution peaks of monomer and hexamer by the calibration curve is indicated as dashed lines. Under the condition of 0.2 M NaCl, the rLonC domain was not eluted because of adsorption to the filter or the resin.

DNA-binding activity of rLonC is essential for the DNA repair function of RadA in vivo

One of the most important questions is whether the DNA-binding activity of rLonC is necessary for the DNA repair function of RadA in vivo. The previous results have shown that rLonC domain-truncated mutants cannot complement the phenotypes of disruptants of the radA genes in E. coli or D. radiodurans (16, 22). However, there is no evidence that the DNA-binding activity of rLonC is important for the DNA repair function of RadA. To address the effect of DNA-binding activity of rLonC on RadA-mediated DNA repair in vivo, we performed an in vivo complementation assay using two DNA-binding-deficient mutants, namely R286A/R385A and R305A/R314A/K345A, in T. thermophilus HB8.

Furthermore, we constructed R286A/R385A and R305A/R314A/K345A mutants because each single mutation of these residues exhibited significant but only limited effects on DNA-binding activity. As a result, both R286A/R385A and R305A/R314A/K345A mutants exhibited lower DNA-binding activity than each single mutant (Fig. 4B and Table 2). The dissociation constants of R286A, R305A, R314A, K345A, R385A, R392A, R314A/K345A, and R385A exhibited slightly decreased DNA-binding activities compared with the wild type, whereas the others exhibited DNA-binding activities similar to the wild type (Fig. 4B and Table 2). The dissociation constants of R286A, R305A, R314A, K345A, and R385A mutated were 1.9–3.4-fold higher than that of the wild type. Two residues, Arg-286 and Arg-385, and three residues, Arg-305, Arg-314, and Lys-345, were located at the unique positively charged regions, namely the intersubunit cleft and the central hole, respectively (Figs. 1E and 4A). The other six residues were located at the outer wall and peripheral regions of the ring. These results suggest that both the intersubunit cleft and the central hole participate in DNA binding and that the outer wall and peripheral regions do not. Interestingly, both of the possible DNA-binding regions are located at subunit interface and formed by dimerization, suggesting the requirement of dimerization for DNA binding (Fig. 4A).

Multiple sequence alignment shows that these possible DNA-binding residues, except for Arg-314 (replaced with Asn in most organisms), are highly conserved in RadA orthologs (Fig. 1C, cyan and yellow backgrounds). By contrast, these basic residues are not conserved in Lon proteases, suggesting that the conservation of these residues is a unique feature of RadA. This finding supports our results that both the intersubunit cleft containing Arg-286 and Arg-385 and the central hole containing Arg-305, Arg-314, and Lys-345 are important for the DNA-binding activity of RadA via rLonC.

The previous results have shown that rLonC domain-truncated mutants cannot complement the phenotypes of disruptants of the radA genes in E. coli or D. radiodurans (16, 22). However, there is no evidence that the DNA-binding activity of rLonC is necessary for the DNA repair function of RadA. To address the effect of DNA-binding activity of rLonC on RadA-mediated DNA repair in vivo, we performed an in vivo complementation assay using two DNA-binding-deficient mutants, namely R286A/R385A and R305A/R314A/K345A, in T. thermophilus HB8.
First, we constructed a \( \Delta radA \) strain of \( T. \) thermophilus HB8 using a gene-disruption method via natural transformation and HR (Fig. 5, A and B) (40). To verify the involvement of \( RadA \) in DNA repair, the phenotype of \( \Delta radA \) was characterized by measuring the sensitivity to UV light and mitomycin C, which introduce DNA strand breaks via DNA adducts and cross-links (Fig. 5C). \( \Delta radA \) was \( \sim 30 \)-fold more sensitive than the wild type to both UV and mitomycin C treatments. These results imply that \( RadA \) functions as a DNA-repair protein in \( T. \) thermophilus HB8 as it does in other organisms, such as \( E. \) coli and \( D. \) radiodurans (16, 22).

The DNA-damage sensitivities of \( \Delta radA \) were recovered at the same level as those of wild type by introducing the expression plasmid of wild-type \( RadA \) (Fig. 5C). This indicates that the \( RadA \) protein expressed by the plasmid was fully functional \( \textit{in vivo} \) and that the DNA-damage sensitivities of \( \Delta radA \) were caused by its loss of \( RadA \). In contrast, overexpression of wild-type \( RadA \) in the wild-type strain showed no effect on sensitivity to UV light and to mitomycin C (Fig. 5, B and C). It should be noted that in the Western blotting analysis, the bands of degraded \( RadA \) were observed even in the presence of protease inhibitors (Fig. 5B). The His tag-cleaved \( RadA \) (labeled with

\[ \text{Figure 3. DNA-binding activity of RadA domain-truncated mutants.} \]

\[ A \text{ and } B, \text{ DNA-binding activity of the domain-truncated mutants of RadA was assessed by EMSA. The representative gels are shown in } A. \text{ The protein concentrations of wild-type (WT) RadA and RadA mutants were 1.1, 2.0, 3.5, 6.0, 11, and 20 } \mu \text{M. Schematic representations of the domain organization of each mutant are shown above the panel as follows: light blue, a ZF domain (Z); pink, a RecA-like ATPase domain (R); and red, a LonC domain (L). The quantitative results of the protein concentration dependence in DNA binding are shown in } B. \text{ The WT and mutants are colored as follows: black, WT; orange, } \Delta \text{N53; pink, } \Delta \text{N53/260C; red, } \Delta \text{N261; green, } \Delta 260C; \text{ and light blue, } \Delta 54C. \text{ The error bars indicate the standard deviation of the values of fraction DNA bound from three independent experiments.} \]

\[ C, \text{ structural comparison of the RecA-like ATPase domain of RadA from } T. \text{ thermophilus HB8 (residues 54–260) with the RecA-dsDNA complex (top; PDB code, 3CMX) and RecA-ssDNA complex (bottom; PDB code, 3CMW) from } E. \text{ coli. The structural model of the ATPase domain of RadA was created using the I-TASSER server (30). The predicted structure of the ATPase domain of RadA and the structure of that of RecA are aligned (r.m.s. deviation } = 2.7 \text{ Å for Ca atoms) and shown in a schematic representation, colored in yellow and orange, respectively. The L1 and L2 loops of RecA are colored in magenta and cyan, respectively. The L1 and L2 loops of RadA are colored in light pink and pale cyan, respectively.} \]

\[ D, \text{ multiple sequence alignment of the ATPase domain of RecA and RadA. The alignment was generated by Muscle (65). The Walker-type ATPase motifs and the conserved residues that are important for DNA binding are highlighted with blue and red backgrounds, respectively. The L1 and L2 loops are boxed in magenta and cyan, respectively. The accession numbers of the sequences are as follows: } T. \text{ thermophilus } (Tt) \text{ RecA (YP_145084); } D. \text{ radiodurans } (Dr) \text{ RecA (NP_296061); } B. \text{ subtilis } (Bs) \text{ RecA (NP_389576); } E. \text{ coli } (Ec) \text{ RecA (NP_417179); } Tt \text{ RadA (YP_143807); Dr RadA (NP_294829); Bs RadA (NP_387968); and Ec RadA (NP_418806).} \]
**Novel Lon protease-like DNA-binding domain of RecA paralog**

RadA in Fig. 5B) was observed in the strains that were complemented with plasmids encoding the His-tagged RadA.

By contrast, neither the expressions of R286A/R385A nor of R305A/R314A/K345A mutants were able to rescue the DNA-damage sensitivities of ΔradA (Fig. 5C). It should be noted that the expression level of the R286A/R385A mutant in ΔradA was apparently lower than that of the wild type and R305A/R314A/K345A mutant (Fig. 5B), implying the structural instability of the R286A/R385A mutant in vivo. These data suggest that these five residues located at the intersubunit cleft and the central hole are important for the proper function of RadA and that at least the DNA-binding ability of RadA via the central hole of rLonC is essential for RadA-mediated DNA repair in vivo.

**Table 2**
The DNA-binding affinities determined from EMSA

| Mutant          | $K_a^{a}$ | $H_a^{b,c}$ |
|-----------------|-----------|-------------|
| WT              | 1.4 ± 0.1 | 2.6 ± 0.3   |
| Domain truncation | 23 ± 2   | 1.9 ± 0.3   |
| ΔN53/260C       | ~         | ~           |
| ΔN261           | 2.0 ± 0.1 | 3.3 ± 0.3   |
| Δ260C           | ~         | ~           |
| Δ54C            | ~         | ~           |
| Site-directed mutagenesis |       |             |
| R286A            | 3.0 ± 0.1 | 3.0 ± 0.3   |
| R305A            | 2.7 ± 0.1 | 2.9 ± 0.4   |
| R314A            | 2.8 ± 0.1 | 3.3 ± 0.4   |
| R345A            | 3.0 ± 0.2 | 2.7 ± 0.4   |
| R385A            | 4.7 ± 0.2 | 2.7 ± 0.3   |
| R392A            | 1.5 ± 0.1 | 2.8 ± 0.3   |
| R395A            | 1.4 ± 0.1 | 2.1 ± 0.3   |
| R399A            | 1.4 ± 0.0 | 2.4 ± 0.2   |
| R404A            | 1.3 ± 0.0 | 2.6 ± 0.2   |
| K412A/R413A      | 1.3 ± 0.0 | 2.4 ± 0.1   |
| R286A/R385A      | 10 ± 1    | 1.5 ± 0.2   |
| R305A/R314A/K345A | 11 ± 1    | 1.8 ± 0.2   |

* The errors are standard errors from the regression analysis.
* $H_a$ is Hill coefficient.
* – indicates no binding activity was detected.

**Biochemical characterization of the DNA-binding-deficient mutants**

The next question is whether there are any functional differences between the two DNA-binding sites. To address this question, we examined the DNA-dependent ATPase activity of wild-type RadA and the DNA-binding-deficient mutants, namely R286A/R385A and R305A/R314A/K345A. In the case of wild type, the turnover rate was stimulated 4.5- and 2.2-fold in the presence of dsDNA and single-stranded DNA (ssDNA), respectively (Fig. 6A and Table 3). In E. coli, it has been shown that dsDNA and ssDNA stimulate the ATPase activity of RadA at the same level (23). Our data suggest that dsDNA and ssDNA have different actions on the ATPase activity of RadA. By contrast, the Michaelis constant for ATP was not affected by the addition of DNA, and the dissociation constants for dsDNA and ssDNA were comparable (Fig. 6, A and B; Tables 3 and 4).

Interestingly, the ATPase activity of R286A/R385A was not stimulated even at a high concentration of DNA (Fig. 6, A and B;
Novel Lon protease-like DNA-binding domain of RecA paralog

Figure 5. Gene disruption and in vivo complementation assays using the DNA-binding-deficient mutants. A, confirmation of gene disruption by genomic PCR. Agarose gel electrophoresis of the PCR products (left panel) and a schematic representation of the genomic regions around the radA gene of the wild type and radA mutant (right panel) are shown. The primers used in the PCR are indicated by black arrows. The theoretical lengths (bp) of the PCR products and the primer sets are shown. The sequences of the primers are listed in Table 5. B, estimation of protein expression levels in T. thermophilus HB8 cells by Western blot analysis using anti-RadA and anti-SSB antisera. C, effects of DNA-binding deficient mutations to UV light (left panel) and mitomycin C (right panel) sensitivity in T. thermophilus HB8 cells. The sensitivity is shown as log10-transformed survival fraction. The mean value (black line) of six independent experiments and each of their values (gray circle) are indicated. The error bar indicates the standard error of the mean. Statistical analysis was performed using Welch’s t test, and multiplicity was adjusted using Holm-Bonferroni method: *, p < 0.05; **, p < 0.01; ***, p < 0.001; ****, p < 0.0001. Statistical comparisons evaluated as “not significant” are not indicated. The left panel and right panel show the R286A/R385A and R305A/R314A/K345A mutants, respectively. EV, empty vector; WT, wild type.

Table 3). Furthermore, R286A/R385A exhibited a higher turnover rate in ATP hydrolysis than the wild type in the absence of DNA, and the turnover rate of R286A/R385A was comparable with that of the wild type in the presence of DNA. These data suggest that Arg-286 and Arg-385 have an important role in the DNA-dependent stimulation of ATP hydrolysis and that the R286A/R385A mutant mimics the stimulated state of wild-type RadA induced by DNA binding. In contrast to R286A/R385A, R305A/R314A/K345A showed similar patterns to the wild type in DNA-dependent ATPase activity, except for a slightly lower turnover rate and binding affinity to DNA than the wild type (Fig. 6, A and B; Tables 3 and 4). Based on these results, we concluded that these two DNA-binding sites have different functions in DNA-dependent ATPase activity.

Discussion

In this study, we determined the X-ray crystal structure of rLonC, which represents the first report of the tertiary structure of the “non-protease-type” LonC domain (Fig. 1). We further revealed that rLonC possesses DNA-binding activity, which is exerted by two positively charged regions, namely the intersubunit cleft and central hole of the ring (Figs. 1, 3, and 4). Our data raise the hypothesis for the model of two DNA-binding modes: one is that the DNA molecule binds to the intersubunit cleft and the other is that the DNA molecule passes through the central hole (Fig. 6C).

As for the latter mode, it should be noted that the loop containing TPFAP forms a lid-like structure at the opposite side of the basic side of the ring (Fig. 1D, left panel). The inner diameter between the two loops is ~17 Å. Therefore, this loop could be steric hindrance based on the hypothesis that the DNA molecule passes through the central hole of the ring. However, the average B-factors of the Cα atoms of this loop and the overall structure were 92.0 and 51.2 Å², respectively, implying a high flexibility of this loop. These data raise the hypothesis that structural rearrangements in the loop region might occur upon binding to DNA through the central hole of the ring.

It should be also noted that RadA forms various types of oligomers that are smaller than a hexamer in solution (Fig. 2), implying the possibility that there are DNA-binding modes other than the hexameric ring structure. Both of the DNA-binding sites are located at the subunit interface and formed by dimerization (Fig. 4A). We speculate that dimerization, not hexamerization, is required and sufficient for DNA binding.

Our biochemical analysis suggests the functional difference between the two DNA-binding sites in the DNA-dependent ATPase activity is that the intersubunit cleft is essential to the DNA-dependent stimulation of the ATPase activity, whereas the central hole is not (Fig. 6C). Interestingly, the R286A/R385A mutant exhibited similar ATPase activity to the DNA-bound stimulated state of the wild type even in the absence of DNA. Thus, Arg-286 and/or Arg-385 would be required to regulate ATPase activity in the DNA-free state as well as to stimulate the activity upon DNA binding. We speculate that DNA binding to the intersubunit cleft of rLonC changes the ATPase domain in such a way as to stimulate its ATPase activity.

It should be noted that RadA can bind to DNA in the absence of nucleotides (Fig. 3, A and B). Interestingly, it has been shown that E. coli RadA requires ADP for DNA-binding activity, suggesting the ATPase domain has an important role for DNA binding (23). Further biochemical analysis and structural information about full-length RadA and its DNA complex are necessary for a complete understanding of how the rLonC and ATPase domains coordinate to regulate DNA-binding activity.
Our in vivo study suggests that the DNA-binding activity of rLonC via the central hole is essential to the DNA-repair function of RadA (Fig. 5). The DNA-binding activity via the inter-subunit cleft is likely to be associated with the DNA-binding function of RadA, although it could not be clearly verified that the loss of DNA-binding activity in the mutant at this region directly leads to the loss of DNA repair activity. The HR pathway involves two DNA molecules and produces branched DNA structures as intermediates, such as the D-loop and Holliday junction (1, 4). It has been suggested that the two DNA-binding sites of RecA/Rad51 are necessary for efficient recombining of two DNA molecules. For instance, RecA has a unique C-terminal domain containing conserved basic residues that interact with dsDNA as well as the L1 and L2 loops in the ATPase domain (41). In addition, the N-terminal regions of Rad51, Dmc1, and archaeal RadA (not a bacterial RadA ortholog) contain a modified helix-hairpin-helix motif, which has a similar function to the C-terminal domain of RecA (42–44). In *E. coli*, genetic and biochemical studies have shown that RadA might be involved in recombination intermediate processing in vivo and that RadA stimulates RecA-catalyzed branch migration in vitro (14, 16, 23). Thus, the two DNA-binding sites of rLonC might be necessary for efficient recombination intermediate processing, such as branch migration.

Our in vivo study revealed the novel LonC subfamily, rLonC, that has a DNA-binding activity instead of a protease activity. We speculate that the LonC family can be divided into “DNA-bind-
ing type” and “protease type.” In the Pfam database (45), most of the rLonC and pLonC have a ChlI (PF13541) signature as well as a Lon_C (PF05362) signature. This ChlI signature was originally found in an N-terminal domain of ComM, which is a widely conserved bacterial protein involved in genetic recombination (46–48). Interestingly, like RadA and Lon protease, ComM also has an ATPase domain featuring Walker A and B motifs at the C terminus. Most of the ComM proteins also have the Lon_C signature in their N-terminal ChlI domain. Thus, we speculate that ComM is a novel LonC family protein. Phylogenetic analysis suggests that the LonC family can be divided into three subfamilies: RadA, ComM, and Lon protease (data not shown). The analysis also suggests that RadA and ComM are more closely related to each other than either is to Lon protease. Structural prediction and structural alignment analyses showed that highly conserved Arg and Lys residues are located near Arg-305, Arg-314, and Lys-345, the DNA-binding residues of rLonC (data not shown). Considering the putative function of ComM in natural transformation via HR (46, 47), ComM might be involved in DNA processing via its LonC domain.

It is noteworthy that among three LonC subfamilies, only the Lon protease has the conserved catalytic Ser residue for the protease activity. Interestingly, the LonC family belongs to the ribosomal protein S5 domain 2-like superfamily, which includes various types of nucleic acid-binding proteins, such as ribosomal proteins S5 and S9, DNA mismatch repair protein MutL, DNA topoisomerases, ribonuclease P, and elongation factor G in the SCOP and CATH databases (49, 50). In particular, ribosomal proteins, MutL, and ribonuclease P utilize the domain belonging to this superfamily as a nucleic acid-binding module (51–53). Therefore, we speculate that the ancestor of the LonC family has a nucleic acid-binding role. Structural and biochemical characterization of the LonC domain of ComM will be necessary to better understand the evolutionary relationships of LonC’s subfamilies.

**Experimental procedures**

**Construction of expression plasmids**

An expression plasmid of RadA from *T. thermophilus* HB8 was obtained from the RIKEN BioResource Center (Tsukuba, Japan). The plasmid, pET-11a/c/Tag/His6TEV (clone name TEx05B04), was constructed by inserting the target gene into pET-11a (Novagen, Madison, WI) using Ndel and BamHI. A fragment containing the *radA* gene digested by Ndel and EcoRI was ligated to a pET-HisTEV vector to create pET-HisTEV/ *radA*. The amplified fragments of each protein (residues 1–53, corresponding to the ZF domain), (ENLYFQG). DNA fragments coding for domain-truncated mutants, R286A, R305A, K314A, K345A, R358A, R392A, R395A, R399A, R404A, and K412A/R413A, were constructed by the QuikChange site-directed mutagenesis method using pET-HisTEV/radA as a template (Stratagene, La Jolla, CA). Expression plasmids for R286A/R385A and R305A/R314A/K345A mutants were constructed by repeating the QuikChange mutagenesis process two and three times, respectively. The primer sequences used for plasmid construction are listed in Table 5.

**Protein overexpression and purification**

*E. coli* Rosetta 2(DE3) cells (Novagen) was transformed with pET-HisTEV/radA. The cells were cultured for 15 h at 37 °C in 2 liters of an overexpression medium containing 1.6% glycoprotein (Nihon Pharmaceutical, Tokyo, Japan), 1.0% yeast extract (Nihon Pharmaceutical), 0.5% NaCl, 0.75% glycerol, 0.05% glucose, 0.2% lactose, 100 µg/ml ampicillin, and 34 µg/ml chloramphenicol and buffered with 25 mM sodium/potassium phosphate at pH 7.4.

The cells were harvested by centrifugation and suspended in 40 ml of buffer I (50 mM Tris-HCl (pH 7.5) and 500 mM NaCl) containing 20 mM imidazole and 1 mM phenylmethylsulfonyl fluoride, and they were then lysed by ultrasonication on ice. After centrifugation (8,000 × g) for 1 h at 4 °C, the supernatant was loaded onto 2 ml of TALON Metal Affinity Resin (Clontech) pre-equilibrated with buffer I containing 20 mM imidazole. The resin was washed with 100 ml of buffer I and then eluted with a 20-ml gradient of 20–300 mM imidazole in buffer I. The fraction containing the His6-tagged RadA was dialyzed against buffer I; then the protein was digested by TEV protease for 6 h at 15 °C. After complete cleavage of the His6 tag, the tag-free RadA was precipitated by 30% saturated ammonium sulfate. After centrifugation (8,000 × g) for 30 min at 4 °C, the precipitant was dissolved in buffer I, and then the solution was loaded onto a HiLoad 16/60 Superdex 200 pg column (GE Healthcare, Uppsala, Sweden) with the same buffer. The eluted RadA was concentrated using a Vivaspin concentrator (Sartorius AG, Gottingen, Germany). The protein concentration was determined on the basis of the absorbance at 280 nm using the molar extinction coefficients (m–1 cm–1) calculated through the previously described procedure (54) as follows: wild type and site-directed mutants, 19,005; ΔN53, 10,290; ΔN261, 4,410; Δ54C, 8,715; Δ261C, 14,595; and ΔN53/261C, 5,880. The purity of each protein (>95%) was assessed by SDS-PAGE.

Except for the rLonC used for X-ray crystallography, all of the RadA mutants were overexpressed and purified with the same procedures as mentioned above. For crystallization, rLonC was purified in a His6-tagged form without TEV protease digestion.

**X-ray crystal structure determination**

A crystal used for X-ray crystallographic analysis was obtained after 6 months of incubation under 4 mg/ml His6-tagged rLonC in 50 mM Tris-HCl (pH 7.5) and 500 mM NaCl in a glass test tube at 4 °C. The crystal was soaked in 25% glycerol as a cryoprotectant and was cryo-cooled in a nitrogen cryostream. Data were collected at the wavelength of 1.00000 Å at
Novel Lon protease-like DNA-binding domain of RecA paralog

Table 5
Sequences of primers used for this study

| Domain truncation | Primer sequence |
|-------------------|-----------------|
| ΔN53              | 5’-ATATCCTATCTGCCAGTGGTACAGGATTGCA-3’ |
| ΔN261             | 5’-ATATCCTATCTGCCAGTGGTACAGGATTGCA-3’ |
| Δ54C              | 5’-TGTATAGTTATGGTGCAAGG-3’ |
| Δ261C             | 5’-TAATACACTACATTAGGCG-3’ |

| Site-directed mutagenesis | Primer sequence |
|--------------------------|-----------------|
| R268A                    | 5’-CCCCTCCGCAGGCAGGCGCAGTCGGGACT-3’ |
| R305A                    | 5’-CCCCTCCCAGGGCGGCGGGCGGTTCGGCCG-3’ |
| R314A                    | 5’-TATGCTAGTTATTATTGCAAGG-3’ |
| R345A                    | 5’-TCTCCGGCGGCTGCCAGGCGAGCTGGGAGG-3’ |
| R385A                    | 5’-GCCCAAGCAGGCGGCGGCGGTTCGGCCG-3’ |
| R392A                    | 5’-GCGCCGCCCTCCGCGAAGAAAGCGGTTCGGCCG-3’ |
| R395A                    | 5’-CTGCTCGCAGGCGGTTCGGCCG-3’ |
| R399A                    | 5’-CCTTAACTCGCAGGCGGCGGTTCGGCCG-3’ |
| R404A                    | 5’-GCCCAAGCAGGCGGCGGCGGTTCGGCCG-3’ |
| K412A/R413A             | 5’-CACAGCCTTCGCGGCCAGGCGGTTCGGCCG-3’ |

| Gene disruption | Primer sequence |
|----------------|-----------------|
| 1              | 5’-CAGCGACGATTCAGGCTACACCCGTAC-3’ |
| 2              | 5’-GAGCCGACGATTCAGGCTACACCCGTAC-3’ |
| 3              | 5’-CAGCGACGATTCAGGCTACACCCGTAC-3’ |

The recognition sites of the restriction enzymes used for plasmid construction and the start/terminator codons are indicated by underlining and boldface, respectively.

Gel filtration analysis

Gel-filtration analysis was performed using a Superdex 200 10/30 GL column (GE Healthcare) on an ÄKTA explorer system (GE Healthcare) at 25 °C. First, 5 µM full-length RadA or rLonC was incubated in a buffer composed of 50 mM Tris-HCl (pH 7.5), 0.2–2 M NaCl at 25 °C for 2 h. Then, 100 µl of the protein solution was loaded on the column and eluted at a flow rate of 0.5 ml/min in the same buffer. The elution profile was monitored by recording the absorbance at 230 nm. The apparent molecular mass was estimated by the calibration curve using thyroglobulin (669,000 Da), β-amylase (200,000 Da), alcohol dehydrogenase (150,000 Da), bovine serum albumin (66,000 Da), ovalbumin (44,000 Da), and carbonic anhydrase (29,000 Da).

EMSA

Chemically synthesized 60-mer ssDNA (Bex, Tokyo, Japan), 5’-GGGTGAAACCTCGAGTTGGCCAAAAGATGTCTTAGCTAGATACTCAACTTCAGTACAGGATTTGCTAGGACATCTTTCGACGCCCTGAGGTCCATTTACACACCAATTTAGGATCAATCTTCTCA-3’, was annealed to complementary 60-mer ssDNA (Bex), 5’-CCGCTACAGTGATCACTCAATAGGTATGGGATAGACATCTTTCGACGCCCTGAGGTCCATTTACACACCAATTTAGGATCAATCTTCTCA-3’, to obtain 60-bp dsDNA. Then, 3 µM (bp) dsDNA and 1.1 to 20 µM RadA were incubated in a buffer containing 50 mM Tris-HCl (pH 7.5), 100 mM NaCl, 0.1 mg/ml bovine serum albumin, 10% glycerol, and 0.01% bromphenol blue at 25 °C for 30 min. The reaction mixtures (5 µl) were loaded onto a 5% polyacrylamide gel containing 10% glycerol and then electrophoresed in 1 × TBE buffer (89 mM Tris borate and 2 mM EDTA). The gel was stained with SYBR Gold Nucleic Acid Gel Stain (Molecular Probes, Eugene, OR), and the bands were visualized by a FLA-3000 image analyzer (Fujifilm, Tokyo, Japan). The amounts of free and shifted DNAs were quantified using MultiGauge version 2.1 software (Fujifilm) to determine the fraction bound. Under conditions with a large excess of RadA over DNA, the apparent dissociation constant $K_d$ was determined from three independent experiments by fitting to Equation 1 using the Igor Pro 4.03 software (WaveMetrics, Lake Oswego, OR).

$$\text{fraction DNA-bound} = \frac{[P]^n}{K_d^n + [P]^n} \quad (\text{Eq. 1})$$

$[P]$, $K_d$, and $h$ are the total protein concentration, apparent dissociation constant for DNA, and Hill coefficient, respectively.
ATPase assay

To determine the kinetic constants of the ATPase activity, an ATPase assay was performed in reaction mixtures containing 50 mM Tris-HCl (pH 7.5), 150 mM NaCl, 10 mM MgCl₂, 0.1 mg/ml bovine serum albumin, 0.5 µM RadA, and 7.8 to 500 µM ATP containing 0.1 µCi of [γ-³²P]ATP, in the presence and absence of 30 µM (bp) dsDNA (60 bp), or 30 µM nucleotides of ssDNA (60-mer). Reactions were initiated by ATP and carried out at 25 °C. To stop the reactions, aliquots of the mixtures were spotted onto a polyethyleneimine cellulose plastic sheet (Merck Millipore, Darmstadt, Germany) and analyzed by thin-layer chromatography with a development buffer containing 0.1 M formic acid and 0.2 M LiCl at room temperature for 5 min. The plate was dried, placed in contact with an imaging plate, and then analyzed by a FLA-3000 image analyzer. The amounts of ATP and Pi were quantified using MultiGauge version 2.1 software. The apparent rate constant \( k_{\text{app}} = \frac{v_0}{[E]_0} \) was estimated from the time-course analysis. The kinetic parameters \( k_{\text{cat}} \) and \( K_m \) were determined from three independent experiments by fitting to Equation 2 using the Igor Pro 4.03 software.

\[
k_{\text{app}} = \frac{k_{\text{cat}}[\text{ATP}]/(K_m + [\text{ATP}])}{[\text{DNA}]/(K_d - [\text{DNA}])}
\]

(Eq. 2)

\( k_{\text{app}} \), \([E]_0\), [ATP], \( k_{\text{cat}} \), and \( K_m \) are the initial velocity, total enzyme concentration, free ATP concentration, catalytic rate constant, and Michaelis constant, respectively.

To determine the dissociation constants for DNA through the DNA-dependent ATPase activity, an ATPase assay was performed in reaction mixtures containing 50 mM Tris-HCl (pH 7.5), 150 mM NaCl, 10 mM MgCl₂, 0.1 mg/ml bovine serum albumin, 0.1 µM RadA, 500 µM ATP containing 0.1 µCi of [γ-³²P]ATP, 0–30 µM (bp) dsDNA (60 bp) or 0–30 µM (nucleotides) ssDNA (60-mer). The initiation and termination of the reactions and the analysis by thin-layer chromatography were performed in the same manner as described above. The apparent rate constant \( k_{\text{app}} \) was estimated from the time-course analysis. The apparent rate constant of the DNA-free enzyme, \( k_{\text{unbound}} \), was determined from six independent experiments in the DNA-free condition. Under conditions with a large excess of DNA over RadA, the kinetic parameters \( k_{\text{bound}} \) and \( K_d \) were determined from three independent experiments by fitting to Equation 3 using the Igor Pro 4.03 software.

\[
k_{\text{app}} = k_{\text{unbound}} + (k_{\text{bound}} - k_{\text{unbound}}) \times [\text{DNA}]/(K_d - [\text{DNA}])
\]

(Eq. 3)

[DNA], \( k_{\text{bound}} \), and \( K_d \) are the total DNA concentration, apparent rate constant of the DNA-bound enzyme, and dissociation constant for DNA, respectively.

Culture conditions for T. thermophilus HB8

T. thermophilus HB8 was grown at 70 °C in TR medium: 0.4% tryptone (Difco, Detroit, MI), 0.2% yeast extract (Nihon Pharmaceutical), and 0.1% NaCl (pH 7.4, adjusted with NaOH). To make TT plates, 1.5 mM CaCl₂, 1.5 mM MgCl₂, and 1.5% gellan gum (Nacalai Tesque, Kyoto, Japan) were added to TR medium.

Constructs for gene disruption and complementation

A T. thermophilus HB8 strain lacking radA (ΔradA) was constructed by substituting the target gene with a thermostable kanamycin-resistance gene, htk, through natural transformation and HR as described previously (40). A plasmid used for gene disruption was obtained from RIKEN BioResource Center. The plasmid (clone name TD03G08) is a derivative of pGEM-T Easy (Promega, Madison, WI) and was constructed by inserting htk flanked by 500 bp upstream and downstream sequences of the target gene. Gene disruption was confirmed by PCR using the isolated genomic DNA as a template and Western blotting using cell lysates. The primer sequences used to confirm gene disruption are listed in Table 5.

For in vivo complementation analysis, gene expression of radA derivatives in T. thermophilus HB8 was carried out using a polycistronic transcription of the radA gene following a thermostable hygromycin B resistance gene under the isocitric acid dehydrogenase promoter from Thermus aquaticus YT1 as described previously (64), with a slight modification. The pMK18::HygΔKm plasmid was constructed by self-ligation after inverse PCR to delete the htk region of pMK18::Hyg. The pET-HisTEV/radA, pET-HisTEV/radA_R286A/R385A, and pET-HisTEV/radA_R286A/R314A/K345A plasmids were digested by XbaI and HindIII and then ligated into the complement site of the pMK18::HygΔKm plasmid to generate pMK18::HygΔKm::radA, pMK18::HygΔKm::radA_R286A/R385A, and pMK18::HygΔKm::radA_R305A/R314A/K345A. These plasmids and pMK18::HygΔKm, as a negative control, were transformed into the T. thermophilus HB8 wild-type and ΔradA strains. Protein expression in each strain was confirmed by Western blotting.

Examination of the sensivities to DNA-damaging agents

The sensitivity to DNA-damaging agents was measured as follows. To measure UV sensitivity, T. thermophilus HB8 cells in the mid-exponential growth phase were serially diluted by TR medium; then the cells were spread onto TT plates and irradiated with 254 nm UV light at a dose of 54 J m⁻² using a SLUV-4 lamp (254 nm; As One, Osaka, Japan). The numbers of colonies were counted after incubation at 70 °C for 18 h.

For excision repair experiments, cells were further incubated at 37 °C for 30 min and spread onto TT plates. The numbers of colonies were counted after incubation at 70 °C for 18 h.

A 12-bp DNA template was constructed by inserting 500 bp upstream and downstream sequences of the target gene. Gene disruption was confirmed by PCR using the isolated genomic DNA as a template and Western blotting using cell lysates. Statistical analysis was performed on log10-transformed data of the surviving fractions from six independent experiments using Welch's t test, and the multiplicity was adjusted by the Holm-Bonferroni method.
Novel Lon protease-like DNA-binding domain of RecA paralog

Western blotting

T. thermophilus HB8 cells at the late-exponential growth phase were harvested by centrifugation. The cells were lysed by ultrasonication in a buffer containing 50 mM Tris-HCl (pH 7.5), 500 mM NaCl, 10 mM EDTA, and 10× protease inhibitor cocktail (Nacalai Tesque). The cell debris was removed by centrifugation. Total protein concentrations in the cell lysates were measured and corrected by the Bradford method using the protein assay kit from Bio-Rad. The cell lysates containing 50 μg of total protein in each lane were resolved on 11% polyacrylamide gels and electroblotted onto a PVDF membrane (Millipore, Milford, MA). The membrane was incubated in a blocking solution containing 20 mM Tris-HCl (pH 7.5), 500 mM NaCl, and 5% skim milk for 30 min at room temperature. After washing with a wash buffer containing 20 mM Tris-HCl (pH 7.5) and 500 mM NaCl, the membrane was immersed into a blocking solution containing rabbit anti-RadA and anti-SSB antisera and then incubated for 2 h at room temperature. After washing with the wash buffer, the membrane was reacted with goat anti-rabbit IgG-AP conjugate (Bio-Rad) in the blocking solution for 1 h at room temperature. The membrane was washed twice with the wash buffer and then reacted with AP color reagents in a wash buffer containing 20 mM Tris-HCl (pH 7.5) and 500 mM NaCl, the membrane was immersed into a blocking solution containing rabbit anti-RadA and anti-SSB antisera and then incubated for 2 h at room temperature. After washing with the wash buffer, the membrane was reacted with goat anti-rabbit IgG-AP conjugate (Bio-Rad) in the blocking solution for 1 h at room temperature. The membrane was washed twice with the wash buffer and then reacted with AP color reagents in a development buffer (Bio-Rad) for 15 min at room temperature. The color development was stopped by washing with deionized water.

Phylogenetic analysis

A multiple sequence alignment was prepared by the Muscle program (65) using 64 LonC domain sequences consisting of 26 Lon protease, 21 RadA, and 17 ComM proteins. Then, a phylogenetic tree was created with the Mega program (66) using the neighbor-joining method (67).

Author contributions—M. I., K. F., R. M., and S. K. conceived and designed the work. M. I. constructed expression plasmids, prepared proteins, crystallized the protein, solved the X-ray structure, and designed the work. M. I. and R. M. wrote the paper with the help from K. F., N. N., T. Y., and S. K.

Acknowledgments—We express our great thanks to Drs. Genji Kurisu, Hisao Masukata, Kenji Matsuno, and Tetsuro Yonesaki for their critical advice. The crystallographic experiments were performed at BL38B1 in SPring-8 (Proposal 2014B1148). We thank the beamline staff for technical help during data collection.

References
1. Cromie, G. A., Connelly, J. C., and Leach, D. R. (2001) Recombination at double-strand breaks and DNA ends: conserved mechanisms from phage to humans. Mol. Cell 8, 1163–1174
2. Krejci, L., Almaanova, V., Spirek, M., and Zhao, X. (2012) Homologous recombination and its regulation. Nucleic Acids Res. 40, 5795–5818
3. Shin, D. S., Chahan, C., Huffman, J. L., and Tainer, J. A. (2004) Structure and function of the double-strand break repair machinery. DNA Repair 3, 863–873
4. West, S. C. (2003) Molecular views of recombination proteins and their control. Nat. Rev. Mol. Cell Biol. 4, 435–445
5. Brigulla, M., and Wackernagel, W. (2010) Molecular aspects of gene transfer and foreign DNA acquisition in prokaryotes with regard to safety issues. Appl. Microbiol. Biotechnol. 86, 1027–1041
6. Thomas, C. M., and Nielsen, K. M. (2005) Mechanisms of, and barriers to, horizontal gene transfer between bacteria. Nat. Rev. Microbiol. 3, 711–721
7. Bell, J. C., and Kowalczykowski, S. C. (2016) RecA: Regulation and mechanism of a molecular search engine. Trends Biochem. Sci. 41, 491–507
8. Chen, Z., Yang, H., and Pavletich, N. P. (2008) Mechanism of homologous recombination from the RecA-ssDNA/dsDNA structures. Nature 453, 489–494
9. Cox, M. M. (2007) Regulation of bacterial RecA protein function. Crit. Rev. Biochem. Mol. Biol. 42, 41–63
10. Lin, Z., Kong, H., Nei, M., and Ma, H. (2006) Origins and evolution of the recA/RAD51 gene family: evidence for ancient gene duplication and endosymbiotic gene transfer. Proc. Natl. Acad. Sci. U.S.A. 103, 10328–10333
11. Taylor, M. R., Spirek, M., Chaurasiya, K. R., Ward, J. D., Carzaniga, R., Yu, X., Egelman, E. H., Collinson, I. M., Rueda, D., Krejci, L., and Boulton, S. J. (2015) Rad51 paralogs remodel pre-synaptic Rad51 filaments to stimulate homologous recombination. Cell 162, 271–286
12. Suwaki, N., Klare, K., and Tarsounas, M. (2011) RAD51 paralogs: roles in DNA damage signalling, recombinational repair and tumorigenesis. Semin. Cell Dev. Biol. 22, 898–905
13. Thacker, J. (2005) The RAD51 gene family, genetic instability and cancer. Cancer Lett. 219, 125–135
14. Beam, C. E., Saveson, C. J., and Lovett, S. T. (2002) Role for Lon protease, 21 RadA, and 17 ComM proteins. Then, a phylogenetic tree was created with the Mega program (66) using the neighbor-joining method (67).

Author contributions—M. I., K. F., R. M., and S. K. conceived and designed the work. M. I. constructed expression plasmids, prepared proteins, crystallized the protein, solved the X-ray structure, and performed in vivo and in vitro experiments. K. F. performed the data collection for X-ray crystallography. Y. F. performed gene disruption of radA. M. I. and R. M. wrote the paper with the help from K. F., N. N., T. Y., and S. K.

Acknowledgments—We express our great thanks to Drs. Genji Kurisu, Hisao Masukata, Kenji Matsuno, and Tetsuro Yonesaki for their critical advice. The crystallographic experiments were performed at BL38B1 in SPring-8 (Proposal 2014B1148). We thank the beamline staff for technical help during data collection.
26. Rotanov, T. V., Melnikov, E. E., Khaltatova, A. G., Makhovskaya, O. V., Botas, I., Wlodawer, A., and Gustchina, A. (2004) Classification of ATP-dependent proteases Lon and comparison of the active sites of their proteolytic domains. *Eur. J. Biochem.* 271, 4865–4871

27. Pinti, M., Gibellini, L., Nasi, M., De Biasi, S., Bortolotti, C. A., Lannone, A., and Cossarizza, A. (2016) Emerging role of Lon protease as a master regulator of mitochondrial functions. *Biochim. Biophys. Acta* 1857, 1300–1306

28. Van Melderen, L., and Aerts, A. (2009) Regulation and quality control by Lon-dependent proteolysis. *Res. Microbiol.* 160, 645–651

29. Biasini, M., Bienert, S., Waterhouse, A., Arnold, K., Studer, G., Schmidt, T., Kiefer, F., Gallo Cassarino, T., Bertoni, M., Bordoli, L., and Schwede, T. (2014) SWISS-MODEL: modelling protein tertiary and quaternary structure using evolutionary information. *Nucleic Acids Res.* 42, W252–W258

30. Yang, J., Yan, R., Roy, A., Xu, D., Poisson, J., and Zhang, Y. (2015) The I-TASSER Suite: protein structure and function prediction. *Nat. Methods* 12, 7–17

31. Krissinel, E., and Henrick, K. (2007) Inference of macromolecular assemblies from crystalline state. *J. Mol. Biol.* 372, 747–779

32. Lee, A. Y., Hsu, C. H., and Wu, S. H. (2004) Functional domains of *Brevibacillus thermoruber* Lon protease for oligomerization and DNA binding: role of N-terminal sensor and substrate discrimination domains. *J. Biol. Chem.* 279, 34903–34912

33. Rasulova, F. S., Dergousova, N. I., Starkova, N. N., Melnikov, E. E., Rumsh, L. D., Ginodman, L. M., and Rotanova, T. V. (1998) The isolated proteolytic domain of *Listeria D.* Lon protease for oligomerization and DNA binding: clues from crystalline state. *J. Mol. Biol.* 120, E2002–E2008

34. Vagin, A., and Teplyakov, A. (1997) MOLREP — some. *Acta Crystallogr.* D Biol. Crystallogr. 53, 465–470

35. Laskowski, R. A., MacArthur, M. W., Moss, D. S., and Thornton, J. M. (2003) PROCHECK: a program to check the stereochemical quality of macromolecular crystal structures. *Acta Crystallogr. D Biol. Crystallogr.* 60, 2145–2157

36. Laskowski, R. A., MacArthur, M. W., Moss, D. S., and Thornton, J. M. (2000) Differences in LexA regulon structure among Proteobacteria towards a more sustainable future. *J. Mol. Biol.* 306, 746–748

37. Emsley, P., Lohkamp, B., Scott, W. G., and Cowtan, K. (2010) Features and some. *Acta Crystallogr.* D Biol. Crystallogr. 66, 2132–2138

38. Murshudov, G. N., Skubak, P., Vagin, A., and Dodson, E. J. (2011) Overview of the CCP4 suite and current developments. *Acta Crystallogr.* D Biol. Crystallogr. 67, 235–242

39. Ramakrishnan, V. (2002) Crystal structure of the 30 S ribosomal subunit from *Thermus thermophilus*: structure of the proteins and their interactions with 16S rRNA. *J. Mol. Biol.* 326, 725–768

40. Yokoyama, S., Dey, A., and Shibata, T. (1995) The central aromatic residue in loop L2 of RecA interacts with charged residues in the L1 and N-terminal DNA binding motifs of archaeal RadA protein. *PLoS ONE* 2, e858

41. Kinebuchi, T., Kagawa, W., Kurumizaka, H., and Yokoyama, S. (2005) Role of the N-terminal domain of the human DMC1 protein in octamer formation and DNA binding. *J. Biol. Chem.* 280, 28382–28387

42. Finn, R. D., Coggill, P., Eberhardt, R. Y., Eddy, S. R., Mistry, J., Mitchell, A. L., Potter, S. C., Punter, M., Qureshi, M., Sangrader-Vegas, A., Salazar, G. A., Tate, I., and Bateman, A. (2016) The Pfam protein families database: towards a more sustainable future. *Nucleic Acids Res.* 44, D279–D285

43. Finn, T. J., Chothia, C., and Beatty, J. T. (2015) Homologues of genetic transformation DNA import genes are required for *Rhodobacter capsulatus* gene transfer agent recipient capability regulated by the response regulator CtrA. *J. Bacteriol.* 197, 2653–2663

44. Brimacombe, C. A., Ding, H., Johnson, J. A., and Beatty, J. T. (1998) A new transformation-deficient mutant of *Haemophilus influenzae* Rd with normal DNA uptake. *J. Bacteriol.* 180, 746–748

45. Laskowski, R. A., Moss, D. S., and Thornton, J. M. (2000) Differences in LexA regulon structure among Proteobacteria through in vivo assisted comparative genomics. *Nucleic Acids Res.* 32, 6617–6626

46. Andreeva, A., Howorth, D., Chandonia, J. M., Brenner, S. E., Hubbard, T. J., Chothia, C., and Murzin, A. G. (2008) Data growth and its impact on the SCOP database: new developments. *Nucleic Acids Res.* 36, D419–D425

47. Vagin, A., and Teplyakov, A. (1997) MOLREP — some. *Acta Crystallogr.* D Biol. Crystallogr. 53, 465–470

48. Murshudov, G. N., Skubak, P., Vagin, A., and Dodson, E. J. (2011) Overview of the CCP4 suite and current developments. *Acta Crystallogr.* D Biol. Crystallogr. 67, 235–242

49. Ramakrishnan, V. (2002) Crystal structure of the 30 S ribosomal subunit from *Thermus thermophilus*: structure of the proteins and their interactions with 16S rRNA. *J. Mol. Biol.* 326, 725–768

50. Yokoyama, S., Dey, A., and Shibata, T. (1995) The central aromatic residue in loop L2 of RecA interacts with charged residues in the L1 and N-terminal DNA binding motifs of archaeal RadA protein. *PLoS ONE* 2, e858

51. Laskowski, R. A., MacArthur, M. W., Moss, D. S., and Thornton, J. M. (1993) PROCHECK: a program to check the stereochemical quality of protein structures. *J. Appl. Crystallogr.* 26, 283–291

52. Baker, N. A., Sept, D., Joseph, S., Holst, M. J., and McCammon, J. A. (2001) Electrostatics of nanosystems: application to microtubules and the ribosome. *Proc. Natl. Acad. Sci. U.S.A.* 98, 10037–10041
Novel Lon protease-like DNA-binding domain of RecA paralog

63. Dolinsky, T. J., Czodrowski, P., Li, H., Nielsen, J. E., Jensen, J. H., Klebe, G., and Baker, N. A. (2007) PDB2PQR: expanding and upgrading automated preparation of biomolecular structures for molecular simulations. *Nucleic Acids Res.* 35, W522–W525

64. Fukui, K., Nishida, M., Nakagawa, N., Masui, R., and Kuramitsu, S. (2008) Bound nucleotide controls the endonuclease activity of mismatch repair enzyme MutL. *J. Biol. Chem.* 283, 12136–12145

65. Edgar, R. C. (2004) MUSCLE: multiple sequence alignment with high accuracy and high throughput. *Nucleic Acids Res.* 32, 1792–1797

66. Kumar, S., Stecher, G., and Tamura, K. (2016) MEGA7: molecular evolutionary genetics analysis version 7.0 for bigger datasets. *Mol. Biol. Evol.* 33, 1870–1874

67. Saitou, N., and Nei, M. (1987) The neighbor-joining method: a new method for reconstructing phylogenetic trees. *Mol. Biol. Evol.* 4, 406–425