C-terminal region of bacterial Ku controls DNA bridging, DNA threading and recruitment of DNA ligase D for double strand breaks repair.

Stephen Mcgovern, Sonia Baconnais, Pierre Roblin, Pierre Nicolas, Pascal Drevet, Héloïse Simonson, Olivier Piétrement, Jean-Baptiste Charbonnier, Eric Le Cam, Philippe Noirot, et al.

Published online 9 March 2016

Nucleic Acids Research, 2016, Vol. 44, No. 10 ***4785–4806***
doi: 10.1093/nar/gkw149

C-terminal region of bacterial Ku controls DNA bridging, DNA threading and recruitment of DNA ligase D for double strand breaks repair

Stephen McGovern[1], Sonia Baconnais[2], Pierre Roblin[3], Pierre Nicolas[4], Pascal Drevet[5], Héloïse Simonson[1], Olivier Piétrement[2], Jean-Baptiste Charbonnier[5], Eric Le Cam[2], Philippe Noirot[1] and François Lecointe[1,*]

[1]Micalis Institute, INRA, AgroParisTech, Université Paris-Saclay, 78350 Jouy-en-Josas, France, [2]UMR 8126, CNRS, Gustave Roussy Université Paris Sud, Université Paris-Saclay, F-94805 Villejuif, France, [3]SOLEIL Synchrotron, F-91192 Gif-sur-Yvette, INRA-URBIA, F-44316 Nantes, France, [4]MaIAGE, INRA, Université Paris-Saclay, 78350 Jouy-en-Josas, France and [5]I2BC, iBiTec-S, CEA Saclay, UMR 9198, F-91191 Gif-sur-Yvette, France

Received July 19, 2015; Revised February 26, 2016; Accepted February 29, 2016

ABSTRACT

Non-homologous end joining is a ligation process repairing DNA double strand breaks in eukaryotes and many prokaryotes. The ring structured eukaryotic Ku binds DNA ends and recruits other factors which can access DNA ends through the threading of Ku inward the DNA, making this protein a key ingredient for the scaffolding of the NHEJ machinery. However, this threading ability seems unevenly conserved among bacterial Ku. As bacterial Ku differ mainly by their C-terminus, we evaluate the role of this region in the loading and the threading abilities of *Bacillus subtilis* Ku and the stimulation of the DNA ligase LigD. We identify two distinct sub-regions: a ubiquitous minimal C-terminal region and a frequent basic C-terminal extension. We show that truncation of one or both of these sub-regions in *Bacillus subtilis* Ku impairs the stimulation of the LigD end joining activity *in vitro*. We further demonstrate that the minimal C-terminus is required for the Ku-LigD interaction, whereas the basic extension controls the threading and DNA bridging abilities of Ku. We propose that the Ku basic C-terminal extension increases the concentration of Ku near DNA ends, favoring the recruitment of LigD at the break, thanks to the minimal C-terminal sub-region.

INTRODUCTION

The genome of each cell is constantly exposed to a variety of damages caused by endogenous or exogenous factors (1). DNA double strand break (DSB) is a lethal DNA lesion as a single unrepaired DSB is sufficient to trigger cell death (2) and DSB misrepair can lead to chromosome rearrangements, a well-recognized cause of cancerogenesis in mammalian cells (3,4). To survive such deleterious event, cells have evolved two main DSB repair pathways: Homologous Recombination (HR) and Non-Homologous End Joining (NHEJ).

The NHEJ pathway is a ligation process involving core and accessory NHEJ factors (5). The core factors are the Ku protein and a specific DNA ligase. In eukaryotes, the Ku protein is a heterodimer formed by the association of the Ku70 and Ku80 subunits; the ligation complex is constituted by XLF/Cernunnos, XRCC4 and DNA ligase IV. Other factors are recruited to form the complete and active eukaryotic NHEJ nucleoprotein complex which can process DNA ends even if broken ends cannot directly be joined by the DNA ligase complex (DNA-PKcs, Artemis, PNK and μ and λ X-family polymerases (5)). In prokaryotes, the core NHEJ factors are composed of the homodimeric Ku and LigD proteins (6–9). LigD displays a ligase activity that is often associated with additional polymerase and/or nuclease activities (8,10,11). Other Ku partners, so far identified in mycobacteria only, come into play when the main NHEJ pathway is partially deficient (12,13) or to achieve yet undefined roles (14,15).

The eukaryotic Ku displays multiple roles in the NHEJ pathway. Ku is the first NHEJ protein recruited to the DNA broken ends. Ku protects DNA ends from nuclease activities and is involved in the regulation of the DSB repair pathway choice in eukaryotes (4). The human Ku displays a 5′-dRP lyase activity (16) showing that Ku participates also in the DNA repair process; this lyase activity was also characterized in two bacterial Ku proteins (17). Finally, Ku proteins are required as a scaffold for the assembly on the DSBs

*To whom correspondence should be addressed. Tel: +33 1 34 65 20 81; Fax: +33 1 34 65 25 21; Email: francois.lecointe@jouy.inra.fr
Present address: Philippe Noirot, Biosciences division, Argonne National Laboratory, Argonne, IL 60439, USA.

of the DNA ligase components as well as other accessory NHEJ factors. In eukaryotes, the recruitment of these proteins is mainly dependent on the N and C-terminal tails of Ku heterodimers (18) which are not conserved in the bacterial Ku proteins (19,20). The N- and C-terminal tails of Ku border a central core domain which, in the human Ku70/80 heterodimer, forms a ring structure able to encircle a double stranded (ds) DNA molecule (21). This particular structure allows multiple Ku proteins to bind a single dsDNA end since already bound Ku proteins can thread inward the DNA molecule when additional Ku proteins enter by the ends (22,23,24). This threading mechanism allows to position at the DNA ends the other NHEJ factors (25) and to displace other proteins bound close to the DSBs such as histones (26). The Ku core domain is similar in sequence and structure (molecular modeling) between human and bacteria (6,19,20,27).

Several bacterial Ku proteins have been purified and characterized so far. They bind preferentially dsDNA ends, interact with their cognate LigD (at least in the presence of DNA) and stimulate LigD repair activities (6–8,27,28). From these observations, a simple model of the NHEJ pathway in bacteria has been proposed: a DSB is recognized and bound by a Ku homodimer at each end, the DSB-bound Ku recruits the LigD protein which seals the DNA ends together (29). While this model accounts for Ku-dependent stimulation of LigD activities on DSBs, it does not include the various DNA binding properties of bacterial Ku proteins whose roles could have an importance in the efficiency of the NHEJ process. For example, Ku_{Msme} (from *Mycobacterium smegmatis*) and Ku_{Bsub} (*Bacillus subtilis*) bind closed circular dsDNA (17,27). These proteins are composed of a Ku core domain associated with a C-terminal tail highly enriched in positively charged amino acids ((30) and Figure 1). The C-terminal domain of Ku_{Msme} is required to bind DNA duplex without free ends, suggesting a direct interaction with circular DNA (27). In addition, Ku_{Mtub} (*Mycobacterium tuberculosis*), which lacks a positively charged C-terminal domain, was proposed to translocate along linear dsDNA molecule after binding DNA ends, a property reminiscent of the threading of eukaryotic Ku proteins (6). Importantly, a Ku_{Msme} truncated for the lysine rich C-terminal domain requires free DNA ends for binding and inward translocation (27). However, the ability of full size Ku_{Msme} for inward translocation from DNA ends remains to be investigated. Ku_{Paer} (*Pseudomonas aeruginosa*), which also displays a basic C-terminal domain, binds DNA ends but does not thread inward a linear 2 700 base pairs (bp) dsDNA molecule and the truncation of this domain did not restore threading ability (7). Therefore, the ability to thread from DNA ends inward may not be conserved in all bacterial Ku proteins, and the role of the basic extended C-terminal domain of Ku in threading ability remains to be established.

In this work, we directly observed threading of bacterial Ku by electron microscopy and we dissected the role of the C-terminal part of Ku_{Bsub} in its loading at DNA ends and threading abilities. An *in silico* analysis delineated two sub-regions of the C-terminal extension. Adjacent to the Ku core domain we found a sequence conserved in all prokaryotic Ku proteins, which we referred to as the minimal Cter. In most bacterial Ku, this minimal Cter is further extended by a variable sequence that tends to be enriched in basic amino acid residues. We refer to this region as the extended Cter. We purified and characterized the biochemical properties of Ku_{Bsub} truncated of the extended Cter or the complete C-terminal region. Both the minimal and the extended Cter are involved in the Ku dependent stimulation of DSB repair by $LigD_{Bsub}$. However, these two sub-regions have also distinct roles. The conserved minimal Cter is required for the interaction between Ku_{Bsub} and $LigD_{Bsub}$. The extended Cter is involved in the interaction with DNA independently on the presence of DNA ends, promotes DNA bridging and regulates the threading of Ku_{Bsub} inward the DNA molecule following DSBs ends binding.

MATERIALS AND METHODS

Bacterial strains and plasmids

Escherichia coli strains used were MC1061 (31) for plasmid constructions and ER2566 from NEB for protein expression. The sequences of all the primers used in this study are available upon request.

The Ku_{Bsub} protein used in this study had the same sequence than the annotated Ku protein from *B. subtilis* 168 (NP_389224.1) from the methionine at position 18 to the serine at position 311 (amino acids positions of the annotated NP_389224.1, see Results) and was tagged by a hexahistidyl-tag (His_6) at its N-terminus. The Ku_{Bsub} protein was produced from an *E. coli* expression-optimized Open Reading Frame (ORF) created by DNA2.0 (Menlo Park, CA, USA) which was derived from the *ykoV* (named hereafter *ku*) ORF in order to decrease mRNA secondary structures and rare codons patches in *E. coli* without changing the amino-acid sequence of the protein. The optimized ORF was flanked by the coding sequence of a His_6-tag followed by a *NdeI* site on its 5′ side, and by a *XhoI* site on its 3′ side. The resulting ORF was cloned under the T7 promoter of the pJ411 vector (DNA2.0) giving the pSMG249 plasmid.

The ORFs coding for KuΔexCter and Ku core mutant proteins (corresponding to the M18 to P277 (KuΔexCter) or M18 to P256 (Ku core) sequence of the NP_389224.1 protein, see Figure 1) were amplified from the pSMG249. The *NdeI-XhoI* fragment of pSMG249 encompassing the *ku* gene was replaced by the PCR products, giving the plasmid pSMG219 and pSMG224 coding for a N-terminal His_6-tagged KuΔexCter and a N-terminal His_6-tagged Ku core proteins, respectively (named hereafter KuΔexCter and Ku core).

The *B. subtilis ykoU* gene coding for the LigD protein was modified by DNA2.0 to optimize its expression in *E. coli*. The modified gene was flanked by a *NdeI* site on the 5′ side and by the coding sequence of a His_6-tag followed by a *XhoI* site on its 3′ side. The cloning of this modified gene in the pJ411, giving the plasmid pSMG217, allowed the expression of the recombinant C-terminal His_6-tagged LigD protein (named hereafter $LigD_{Bsub}$).

The plasmid pSMG240 allowing expression of the $MRGS(H)_6GS$-tagged Green Fluorescent Protein fused to the basic extended Cter sub-region of Ku_{Bsub} (named hereafter GFP-KC, the extended Cter part of Ku correspond-

ing to the sequence from P277 to S311 of the NP_389224.1 annotated protein) was obtained in two steps. Firstly, the corresponding 3′ part of the *ku* gene was amplified from pSMG249, using primers allowing flanking of the PCR product by *ApaI* and *XhoI* sites on the 5′ and 3′ sides, respectively. This 3′ part of *ku* was cloned in frame at the 3′ of the GFP gene of the plasmid pSG1729 (32) cut by *ApaI* and *XhoI*, giving rise to the plasmid pSMG239. Secondly, GFP-KC gene was amplified from pSMG239 using primers allowing its integration between *BamHI* and *XhoI* sites of plasmid pQE9 (Qiagen).

The MRGS(H)$_6$GS-tagged GFP (named hereafter GFP) protein was obtained *via* the expression of the corresponding gene on the pSMG245. This plasmid was constructed directly by amplification of the GFP gene on the plasmid pSG1729 and its integration between the *BamHI* and *XhoI* sites of the pQE9 plasmid.

Comparative sequence analysis

Profile HMMs are considered as one of the most powerful methods to identify remote protein homologues (PMID:9837738). We used profile HMMs to build a list of Ku proteins found in the 2645 complete prokaryotic genomes available in September 2013 as described below.

The first task consisted in collecting from Genbank a non-redundant and diverse set of sequences annotated as corresponding to Ku proteins that could be used to learn the parameters of the profile HMM. We first retrieved from Genbank the 1048 protein sequences corresponding to protein coding genes that have been annotated as belonging to Cog1273. This data set was subjected to an all-against-all comparison with blastp 2.2.26 (PMID:9254694). This allowed to exclude a small subset of 16 highly atypical sequences that displayed hits (E-value cut-off 0.00001) with less than half of the 1048 protein sequences (this included 6 proteins from the eukaryote *Leishmania*, 2 proteins from bacterium *Rhodanobacter fulvus* without similarity and whose annotation was considered as unsure, and 8 protein fragments of length less than 100 amino acids). The level of protein sequence similarity between a pair of proteins was summarized by the raw Blast Score Ratio defined as BSR[P1,P2] = BS[P1,P2] / max{BS[P1,P1],BS[P2,P2]}, where P1 and P2 are two proteins and BS[P1,P2] is the Blast raw score between P1 and P2 (PMID:15634352). Out of the remaining 1032 protein sequences we used a simple greedy algorithm to select a subset of 89 sequences representatives of the whole diversity defined such that any of the 1032 protein sequences has at least one representative in the subset with BSR $\geq$ 0.5 and BSR $<$ 0.5 between proteins of the subset. The 89 representatives plus the *B. subtilis* 168 Ku core domain (taken as a reference, see Figure 1) were subjected to multiple alignment with MUSCLE v3.8.31 (PMID:15034147). This alignment allowed to delineate the Ku domains of the 89 representatives and was used to identify a set of 71 full-length Ku domains for which the residues that aligned with the reference represented at least 80% of the length of this particular Ku domain and the reference Ku domain.

Using HMMER 3.0 (PMID:22039361), the 71 full-length Ku domains realigned with MUSCLE served to build an HMM profile that was used to query the complete genomes. Our final list of 689 Ku proteins consisted of all the hits scoring above 50% of the highest score (E-values were highly significant $<10^{-47}$ for all those hits). The choice of this score cut-off reflected a marked discontinuity in the distribution of the scores: no hits were found between 54% and 23% of the highest score and only 13 hits with E-value $<$ 0.05 were found below the chosen score cut-off.

The 689 Ku proteins identified in complete genomes were aligned with MUSCLE. The Ku core domain as well as the minimal and extended Cter regions were delineated based on this alignment. Isoelectric points (pI) were determined with Bioperl using EMBOSS amino acid pK values (PMID:12368254, PMID:10827456). A tentative phylogenetic tree was obtained from the multiple alignment with the maximum likelihood algorithm implemented in Phyml v3.0 with JTT model of protein evolution and a gamma distribution (approximated with 4 categories of sites) to account for rate heterogeneity (PMID:20525638). The graphical representation of the tree was drawn with R package ape (PMID:14734327).

Purification of the proteins produced in *E. coli*

Ku$_{Bsub}$ and C-terminal truncated mutants. *E. coli* ER2566 cells transformed with pSMG249 (producing Ku$_{Bsub}$), pSMG219 (producing KuΔexCter) or pSMG224 (producing Ku core) were grown at 30°C in 2 liters of LB medium supplemented with 25 mg/ml thymine and 30 μg/ml kanamycin to A_{600nm} = 0.7. Production of proteins was induced with 0.5 mM IPTG for 3 h at 30°C. Cells were harvested by centrifugation and pellets were resuspended in 40 ml lysis buffer (50 mM Tris-HCl pH 8, 0.2 M NaCl, 0.1% Triton, 0.1% Nonidet P40, 0.2 mg/ml lysozyme) and incubated 1 h at 20°C. Cells were broken by a freeze/thaw step in liquid N_2/37°C. All subsequent steps were carried at 4°C. Two milliliters of polyethyleneimine (PEI) 10%, pH 8 was added to precipitate nucleic acids; the mixture was stirred for 1 h and then centrifuged 1 h at 100 000 g. Soluble phase was next precipitated with raising concentrations of ammonium sulfate (AS) by steps of 10% of saturation, and AS pellets were resuspended with 25 ml buffer P_1 (50 mM Tris-HCl, pH 8, 1 M NaCl). Fractions were checked for the presence of the protein of interest by 10% SDS-PAGE. Ku$_{Bsub}$ was mainly found in the pellet obtained at 50% of AS saturation, KuΔexCter and Ku core in the pellet obtained at 40% of AS saturation. The AS pellets were resuspended with 40 mL of buffer P_1 and were loaded onto a Ni^{2+} affinity column (Ni-NTA agarose, QIAGEN) pre-equilibrated in buffer P_1. The Ni-NTA column was washed with the same P_1 buffer then with P_1 buffer supplemented with 20 mM imidazole. Ku$_{Bsub}$ and C-terminal mutants were eluted in buffer P_1 supplemented with 250 mM imidazole and dialysed against buffer S (50 mM Tris-HCl pH 8, 0.15 M NaCl, 5% glycerol) for storage at −80°C before their use in small angle X-ray scattering (SAXS) analyses or against buffer D (50 mM Tris-HCl pH 8, 0.4 M NaCl, 50% glycerol, 1 mM DTT) before their storage at −20°C for all other experiments.

$LigD_{Bsub}$ protein. *E. coli* ER2566 cells transformed with plasmid pSMG217 (producing $LigD_{Bsub}$) were grown at 30°C in 2 liters of LB medium containing 25 mg/ml thymine and 30 μg/ml kanamycin. At $A_{600nm} = 0.7$, culture was shifted at 16°C before induction with 0.5 mM IPTG for 18 h. Cells were harvested by centrifugation and the pellet was resuspended in 40 ml of P_1 buffer and stored at −20°C. All subsequent steps were carried at 4°C. Cells were lysed by sonication and centrifuged at 100 000 g for 1 h. AS was added to the lysis supernatant up to 30% of AS saturation. The AS pellet was resuspended in 40 ml of P_1 buffer and was loaded onto a Ni^{2+} affinity column pre-equilibrated in buffer P_1. The Ni-NTA column was washed with P_1 buffer and then with P_1 supplemented with 20 mM imidazole. The protein was eluted with the same buffer supplemented with 250 mM imidazole and was dialyzed against buffer D. $LigD_{Bsub}$ was diluted with buffer P_0 (50 mM Tris-HCl, pH 8) to reduce the NaCl concentration to 75 mM and loaded onto a 1 ml Hitrap Heparin (GE Healthcare). $LigD_{Bsub}$ was eluted in one step in buffer P_1. The protein was then dialyzed against buffer D prior storage at −20°C.

GFP and GFP-KC proteins. The GFP was produced from plasmid pSMG245 in *E. coli* ER2566 cells grown at 30°C in 2 liters of LB medium containing 25 mg/ml thymine and 100 μg/ml ampicillin to $A_{600nm} = 0.7$ and induced 18 h with 0.5 mM IPTG at 16°C. After centrifugation, harvested cells were resuspended in 40 ml of buffer P_1, lysed by sonication and centrifuged at 100 000 g for 1 h at 4°C. All subsequent steps were carried at this temperature. Supernatant was loaded onto a Ni^{2+} affinity column pre-equilibrated in buffer P_1. The Ni-NTA column was washed with buffer P_2 (50 mM Tris-HCl, pH 8, 2 M NaCl) supplemented with 25 mM imidazole. The protein was eluted with buffer P_2 supplemented with 250 mM imidazole and dialysed against buffer D prior storage at −20°C.

The GFP-KC protein was produced from a freshly pSMG240 transformed colony of *E. coli* ER2566. Its purification was done as described for the GFP, except that a supplemental purification step was done. The dialyzed protein of the Ni-NTA eluate was diluted with buffer P_0 to reduce the NaCl concentration to 50 mM, loaded onto a cation exchange Hitrap SP 1 ml column (GE Healthcare) equilibrated with buffer $P_{0.05}$ (50 mM Tris-HCl, pH 8, 50 mM NaCl), and eluted with a linear 15 ml gradient of NaCl from 50 mM to 1 M in buffer P_0. The GFP-KC obtained was then dialysed against buffer D and stored at −20°C.

All these proteins were obtained at 95% purity and their concentration was determined by UV absorption at λ = 280 nm on a NanoDrop 1000 Spectrophotometer (Thermo Scientific).

SAXS data collections

Ku_{Bsub} and the two C-terminal truncated mutants were analyzed by SAXS experiments to describe their structural organization. All protein solutions were centrifuged for 10 min at 10 000 g prior to X-ray analysis in order to eliminate aggregates and their concentration was measured by UV. Protein samples were prepared at a final concentration of 95 μM, stored at 4°C and then directly used for the experiments.

SAXS experiments were conducted on the SWING beamline at the SOLEIL synchrotron (λ = 1.033 Å). The Aviex charge-coupled device detector was positioned to collect data in the Q-range 0.008–0.5 $Å^{-1}$ ($Q = 4\pi\sin\theta\lambda^{-1}$, where 2θ is the scattering angle). All solutions were mixed in a fixed-temperature (15°C) quartz capillary with a diameter of 1.5 mm and a wall thickness of 10 μm, positioned within a vacuum chamber. Fifty microliters of samples of proteins at 95 μM were injected onto a size-exclusion column (SEC-3 300 Agilent), using an Agilent HPLC system, and eluted directly into the SAXS flow-through capillary cell at a flow rate of 0.2 ml/min (33). The elution buffer consisted of 50 mM Tris-HCl, pH 8, 150 mM NaCl and 2% glycerol. SAXS data were collected continuously, with a frame duration of 1.0 s and a dead time between frames of 0.5 s. Data reduction to absolute units, frame averaging and subtraction were done using FOXTROT, a dedicated home-made application. All subsequent data processing, analysis and modeling steps were carried out with PRIMUS and other programs of the ATSAS suite (34).

Data evaluation. The experimental SAXS data for all samples were linear in a Guinier plot of the low q region, indicating that the proteins did not undergo aggregation. The radius of gyration R_g was derived by the Guinier approximation $I(q) = I(0)\exp(-q^2R_g^2/3)$ for $qR_g < 1.0$ using PRIMUS (35). Interference-free SAXS profiles were estimated by extrapolating the measured scattering curves to infinite dilution. The program GNOM (36) was used to compute the pair-distance distribution functions, $p(r)$. This approach also features the maximum dimension of the macromolecule, D_{max} (Figure 3A and B). The Porod volume were determined with PRIMUS, the molecular weights were estimated with SAXS-MoW (37) and the Porod coefficients were calculated using SCATTER (38).

Ab Initio modeling. The overall shapes of the entire assemblies were restored from the experimental data using the program GASBOR (39). These models were averaged to determine common structural features and to select the most typical shapes using the programs DAMAVER (40) and SUPCOMB (41).

Molecular modeling. Due to the absence of structure of Ku_{Bsub}, the initial model building for a dimeric Ku_{Bsub} was performed using the structure of the complex human Ku heterodimer-DNA (code pdb : 1JEY, (21)) as template. As previously described (19), the sequence of the Ku core domain of the Ku70 subunit has a better homology with the Ku core domain of the Ku_{Bsub} sequence than the Ku80 subunit and the C-terminal domain of Ku70 and Ku_{Bsub} present a strong sequence divergence. We generated a homodimer composed of two Ku70 deleted of the N-terminal (residues 34 to 260) and C-terminal regions (residues 506 to 534) defining a chimera Ku core domain without DNA, keeping the ring structure of the DNA recognition domain (see Supplementary Results, Figures S2 and S3).

To generate a structural model of the Ku_{Bsub} and its derivative mutants compatible with the SAXS data, we per-

formed atomic modeling using DADIMODO (42), a genetic algorithm based rigid-body refinement analysis program. The initial structures were first minimized by keeping fixed the main chain of the rigid domain. The declared flexible parts (C-terminal regions of the homodimeric Ku_{Bsub}) were then submitted to random mutations, whereby phi/psi angles were changed with a maximum amplitude of 45° per step. Continuity of the structure was constrained by subsequent energy minimization. A SAXS χ^2 value was then computed for each eligible structure, using CRYSOL (43) with 50 harmonics. The selection protocol, based on tournament method (42) with progressive increase of the selection, allows the final selection of the best fitting models (Figure 3). To avoid overfitting and to control validity of the models, a χ^2 free metric was also calculated using SCATTER (38).

DNA preparation

The 279 bp linear dsDNA molecules were obtained by PCR amplification of a region of the *B. subtilis* genome. The 1001 bp and 1876 bp dsDNA molecules were amplified from the pZE14 plasmid (44). PCRs were done using 5′-phosphorylated primers and the Phusion polymerase (Thermofisher) in order to generate blunt-ended products. For the DNA bridging assay, the one-end biotinylated 1001 bp dsDNA molecule was prepared using one 5′-biotinylated primer and one 5′-phosphorylated primer. Amplified DNA molecules were purified on agarose gel using the QIAquick PCR purification kit (QIAGEN).

The pUC19 plasmid (2686 bp) was isolated from *E. coli* DH5α cells using Isolating Recombinant Bacmid DNA protocol (Invitrogen) and purified on an ion exchange Resource Q column (GE Healthcare) with a chromatography SMART system (GE Healthcare). For electron microscopy experiments, all DNA molecules were purified on a MiniQ anion exchange column (GE Healthcare) with a chromatography SMART system. The purified DNA was precipitated and resuspended in a 10 mM Tris-HCl, pH 7.5, 1 mM EDTA buffer.

Ligation assay

One hundred nanograms (∼150 fmol) of the 1001 bp linear dsDNA molecules with 5′-phosphorylated ends were preincubated with 9 pmol of Ku_{Bsub} or C-terminal Ku truncated mutants in 6.52 μl of buffer L (50 mM Tris-HCl pH 8, 1 mM DTT, 5 mM $MgCl_2$, 30 mM NaCl and 3% glycerol) for 30 min at 37°C. Serial dilutions of the $LigD_{Bsub}$ protein in buffer L supplemented with 0.39 mM adenosine triphosphate (ATP) were prepared and 11.48 μl of each dilution were added to the reaction (allowing final concentrations of $LigD_{Bsub}$ indicated in the legend of Figure 2). The final volume of each reaction mixture was 18 μl in buffer L supplemented with ATP (0.25 mM final concentration). Ligation reactions were stopped after 2 h at 37°C by adding 2 μl of stop buffer (5 mg/mL proteinase K, 2% SDS and 0.1 M EDTA) and incubating at 50°C for 30 min. A total of 6.7 μl of DNA loading buffer (33% glycerol and 0.25% xylene cyanol) were added, reactions were cooled on ice, centrifuged 1 min at 14 000 g and half of the reaction volume was loaded onto a 0.8% agarose gel in Tris-Borate-EDTA (TBE) buffer. After electrophoresis, DNA was stained by the SYBR Gold reagent (Invitrogen) and the gel was photographed using a Chemidoc apparatus (Bio-Rad Laboratories). Quantification of the unligated DNA substrate was done using the Image Lab v.5.0 software (Bio-Rad Laboratories).

Gel filtration assay

Five nmol of Ku_{Bsub}, KuΔexCter or Ku core proteins, 2.5 nmol of $LigD_{Bsub}$ or mixes of 2.5 nmol of $LigD_{Bsub}$ with 5 nmol of Ku_{Bsub} or C-terminal truncated mutants were incubated in 600 μl of buffer $P_{0.15}$ (50 mM Tris-HCl, pH 8, 150 mM NaCl) for 20 min at 30°C. After 5 min of centrifugation at 13 000 g, the soluble fraction was injected on a Superdex 200 10/300 GL (GE Healthcare) equilibrated with buffer $P_{0.15}$. Twenty two fractions of 500 μl were collected, ranging from approximately 700 kDa (the void volume) to 10 kDa. To compare the elution volumes of the proteins, 5 μl of SDS-loading buffer were added to 20 μl of each of these fractions and loaded onto 10% acrylamide SDS-PAGE gels. After migration, gels were stained with Coomassie blue.

Electrophoretic mobility shift assay (EMSA)

Various amounts of proteins were added to 50 ng of DNA substrate in 20 μl reaction mixtures in buffer L. After incubation at 30°C for 20 min, reactions were mixed with 4 μl of EMSA loading buffer (50% glycerol, 0.04% xylene cyanol) prior loading on a 5% native polyacrylamide gel (29:1, in Figure 6A) or on a 0.7% GTG-agarose (Lonza) gel in TBE buffer (Figure 5 panels A and F), and then migrated at 4°C at 15 V/cm for 2 h. Following electrophoresis, gels were colored with SYBR-Gold and revealed on a Chemidoc apparatus. Quantification of unbound linear dsDNA molecules was done using the Image Lab v.5.0 software (Bio-Rad Laboratories).

Transmission electron microscopy (TEM) analysis

TEM samples preparations were performed by positive staining as previously described (45). All nucleoprotein complexes were formed in 10 mM Tris-HCl pH 8, 50 mM NaCl or 10 mM Tris-HCl pH 8, 200 mM NaCl binding buffer with 0.3 nM DNA substrates and in presence of various concentrations of Ku proteins. Five microliters of DNA–protein reaction were deposited onto a 600 mesh copper grid coated with a thin carbon film, previously activated by glow-discharge in the presence of pentylamin (Merck, France). Grids were washed with aqueous 2% (w/vol) uranyl acetate (Merck, France) and then dried with ashless filter paper (VWR, France). TEM observations were carried out on a Zeiss 912AB transmission electron microscope in filtered crystallographic dark field mode. Electron micrographs were obtained using a ProScan 1024 HSC digital camera and iTEM software (Olympus, Soft Imaging Solutions).

To characterize the different classes of DNA molecules bound by Ku_{Bsub} or by C-terminal mutants, the number of naked DNA molecules and DNA–protein complexes not

trapped into networks or aggregates per μm^2 was determined at different incubation times. This allowed (i) to calculate the frequencies of these untrapped DNA molecules (and consequently frequencies of networks and aggregates, Figure 8D) and (ii) to estimate the proportion of the different classes of untrapped DNA molecules (Figure 8E). The nucleoprotein complexes untrapped into networks or aggregates were divided in four classes: naked DNA, complexes with proteins at one or both DNA ends and complexes with proteins inside the DNA molecule and one, two or without covered ends (see Figure 8E). The number of molecules in each class is counted, normalized at 100 and represented by a histogram.

To compare the DNA threading ability of the proteins, the DNA length covered by each protein and the total contour length are measured for each nucleoprotein complex untrapped into networks using the iTEM software. The percentage of DNA covered by each protein corresponds to the ratio of the sum of DNA length covered by Ku_{Bsub} or truncated mutants at each end over the total length of the DNA molecule. The distribution is represented by box and whisker plots and was statistically analyzed with a Mann–Whitney U test done with Prism5 software (Graphpad software).

DNA bridging assay

This assay was derived from the DNA bridging assay described in (46). Ten microliters of Streptavidin-coated Dynabeads Kilobase BINDER (Invitrogen) were washed three times with 40 μl of reaction buffer (10 mM Tris-HCl, pH 8, 50 mM NaCl and 400 μg/ml BSA) and resuspended in 40 μl of reaction buffer. Two hundred nanograms of the blunt-ended 1001 bp PCR product either 5′-phosphorylated on both ends (named hereafter 1001PP, control experiment) or 5′-biotinylated on one end and 5′-phosphorylated on the other end (1001BP) were incubated for 5 min with the beads. Beads were washed twice with 40 μl of reaction buffer, resuspended in 36 μl of the same buffer and 4 μl of 40 μM Ku_{Bsub}, KuΔexCter or Ku core solutions were added. After 20 min incubation at 30°C, beads were washed twice with 40 μl of reaction buffer and resuspended in 60 μl of the same buffer. At this step, a 20 μl aliquot was analyzed by SDS-PAGE to check fixation of the proteins (named 'M' fraction in Figure 7A).

Next, 200 ng of the blunt-ended 1876 bp PCR product (5′-phosphorylated at both ends) were added to the remaining 40 μl beads suspension and incubated for 5 min at 30°C. Beads were washed twice with 40 μl of the reaction buffer without BSA and then resuspended in 40 μl of the same buffer.

To evaluate the amount of captured 1876 bp fragment, 20 μl of the beads suspension were treated with 2 μl of stop buffer for 30 min at 50°C and then resolved by electrophoresis on a 0.7% TBE agarose gel. DNA was stained by the SYBR Gold reagent and the gel was photographed using a Chemidoc apparatus. The remaining 20 μl of the bead suspension (named 'E' fraction in Figure 7A) were analyzed by SDS-PAGE.

Nuclease protection assays

The 1001 bp linear dsDNA molecule used in the T5 exonuclease assay was the same as those used in the ligation assay described in the Materials and Method. One hundred nanograms of this 5′-phosphorylated blunt-ended molecule was preincubated with Ku_{Bsub} or KuΔexCter proteins in Buffer L (50 mM Tris-HCl pH 8, 1 mM DTT, 5 mM $MgCl_2$, 30 mM NaCl and 3% glycerol) for 30 min at 37°C. A total of 0.5 U of T5 exonuclease (Biolabs) was added to the reaction mixture (final volume: 18 μl) and reactions were incubated for 1 h at 37°C. Reactions were stopped by adding 2 μl of stop buffer (5 mg/mL proteinase K, 2% SDS and 0.1 M EDTA) and incubating at 50°C for 30 min. Then 6.7 μl of loading buffer (33% glycerol and 0.25% xylene cyanol) were added to the reaction. Degradation products corresponding to the half of the reaction were resolved by electrophoresis on a 0.8% agarose gel in TBE buffer. DNA was stained by SYBR Gold and the gel was photographed under UV light.

A HincII endonuclease assay was performed to test the Ku mediated protection of DNA against restriction. We used the same blunt-ended 1001 bp linear dsDNA molecule with 5′ phosphorylated ends. This molecule contains two HincII sites (see Figure 9). The assay was the same as the T5 exonuclease assay described above except that 50 ng of the 1001 bp substrate was used and the digestion reaction were done in buffer L supplemented with 0.1 mg/ml BSA and the incubation time with HincII (4.8 U) was 2 h at 37°C. Digested products present in 20 μl out of the 26.7 μl of the final reaction volume were analyzed. After SYBR Gold staining, the gel (1% agarose in TBE) was photographed using a Chemidoc apparatus. Quantification of digested products was done using the Image Lab v.5.0 software (Bio-Rad Laboratories).

Kinetics of HincII digestion of the 1001 bp dsDNA substrate preincubated with different Ku concentrations was done as described for the HincII endonuclease assay except that 400 ng of the 1001 bp substrate and 3.84 U of HincII were used in a final reaction volume of 144 μl. Samples of 18 μl were extracted at different time of the incubation at 37°C and analyzed as described above.

RESULTS

Prokaryotic Ku core domain is associated with a conserved minimal Cter region frequently extended by a basic tail

To examine the domain architecture of prokaryotic Ku proteins, we searched 2645 complete prokaryotic genomes for genes encoding putative Ku proteins (see Materials and Methods), and we identified 689 *ku* genes. 20.2% of all genomes contain a unique *ku* gene (528 bacterial genomes and only 7 euryarchaeotal genomes) and 3.5% of genomes contain more than one *ku* (details in Supplementary Table S1). For instance, 4 *ku* homologues are found in the *Sinorhizobium meliloti* genome (47), with some bacteria of the *Rhizobium* genus having 6 *ku* genes. The microorganisms bearing a *ku* gene belong predominantly (518/528) to the Proteobacteria, Actinobacteria, Firmicutes and Bacteroidetes bacterial phyla (Figure 1A). However, even in these phyla, *ku* genes were found only in a fraction of the sequenced genomes: 54.7% of the Actinobacteria, 23.6% of the Fir-

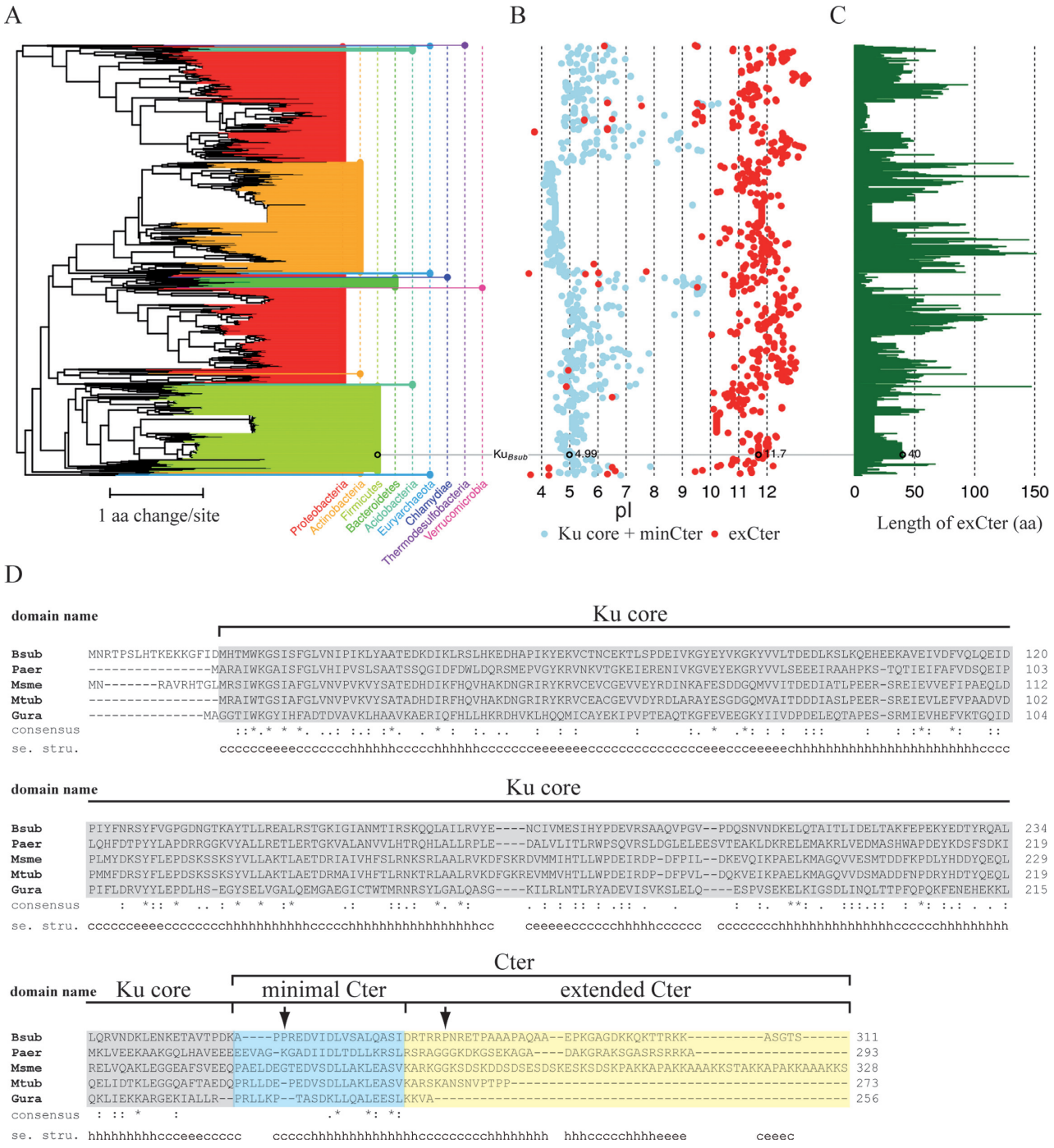


Figure 1. Bacterial Ku proteins contain a minimal C-terminal region which is frequently followed by a basic extended C-terminal domain. (**A**) Phylogenetic analysis of the prokaryotic Ku proteins. The 689 Ku proteins identified in complete prokaryotic genomes are presented at the tips of a tentative phylogenetic tree. Information on the organisms in which each of these Ku encoding genes were identified are provided by horizontal bars of color and length specific of each represented phylum. (**B**) For each of these proteins, the pI of the Ku core and minimal Cter (minCter) is represented (light blue points) along the pI of the extended Cter (exCter, red points). (**C**) The length of the extended Cter is represented by a horizontal green bar. The position of Ku$_{Bsub}$ in the tree is indicated by a horizontal dotted line and the corresponding pI values and length of the extended Cter are reported in plain text. (**D**) Alignments of amino acid sequences of Ku proteins from *Bacillus subtilis* (Bsub, NCBI Reference Sequence: NP_389224.1), *Pseudomonas aeruginosa* (Paer, YP_791093.1), *Mycobacterium smegmatis* (Msme, YP_889815.1), *Mycobacterium tuberculosis* (Mtub, NP_215453.1) and *Geobacter uraniireducens* (Gura, YP_001232184.1). Alignments of amino acid sequences were done using the multiple sequence alignment Clustal Omega method (59) using the default value for each parameter. The conserved Ku core and minimal Cter domains are highlighted in gray and blue respectively. The common extended Cter is highlighted in yellow. Conserved residues between Ku proteins are indicated by an asterisk. Strongly and weakly similar residues are indicated by a column and a period respectively. A hyphen is used to represent a gap in the sequence. The secondary structure prediction (se. stru.) of the Ku$_{Bsub}$ from the first methionine of the Ku core domain to the end of the protein was obtained using the GOR version IV method (60). c, e and h stand for random coil, extended strand and alpha helix, respectively. Arrows indicate last residues of the Ku core (proline pointed in the minimal Cter) and the KuΔexCter mutants (proline pointed in the extended Cter).

micutes, 19.9% of the Proteobacteria and 14.4% of the Bacteroidetes (Supplementary Table S2).

All prokaryotic Ku proteins contain a central Ku core domain ((19), in gray in Figure 1D) displaying high level of homology with the eukaryotic Ku70 proteins and ranging from 187 and 311 amino acid residues (average 236) in length. The Ku core domain is generally very close to the N-terminus as 92.9% of Ku proteins have an N-terminal extension of 5 residues or less (Figure 1D and Supplementary Table S1). Extending the Ku core domain, all prokaryotic Ku proteins display a C-terminal region ranging from 19 to 173 residues (average 59) in length. Sequence alignments delineated two distinct domains in the bacterial Ku C-terminal region. Directly adjacent to the Ku core, a small domain of 12 to 28 residues, hereafter named minimal Cter domain, exhibits some similarity across prokaryotic Ku proteins (Figure 1D). The minimal Cter domain is followed by an extended C-terminal domain (hereafter named extended Cter) in the vast majority of prokaryotic Ku proteins (Figure 1C). The extended Cter domain varies considerably in length, ranging from 1 to 155 residues. A remarkable feature of this domain is that it tends to be markedly basic (Figure 1B), with a mean isoelectric point of 11.4 relative to 5.5 for the Ku core and the minimal Cter domains. This is due to a bias in the amino acid composition (lysine and to a lesser extent arginine residues), already noticed for Ku_{Paer} and Ku_{Msme} (7,30). The extended Cter regions do not share any significant sequence similarity (Figure 1D).

Purification of *B. subtilis* Ku protein variants of the C-terminal region

The *B. subtilis* Ku protein (NP_389224.1) is composed of a 236 residues long Ku core domain with an unusually long (17 residues) N-terminal domain, a minimal Cter domain (18 residues) and an extended Cter domain (40 residues) comprising 11 lysine and arginine residues (isoelectric point 11.7, Figure 1D). The transcription start site (TSS) of the *ykoV* gene encoding Ku_{Bsub} was mapped in a large scale structural reannotation of the *B. subtilis* genome (48). The *ykoUV* operon, which is under the control of the SigG transcription factor, is located downstream of the predicted start codon in UniprotKB, indicating an annotation error. An ATG located 17 codons downstream and corresponding to the first position of the Ku core domain is the likely Ku start site as it is compatible with the mapped TSS (U1083) and consistent with conserved Ku core domain (Figure 1D).

We cloned, expressed in *E. coli* and purified $LigD_{Bsub}$, Ku_{Bsub} and two C-terminally truncated Ku mutants. Ku_{Bsub} was truncated at two sites (indicated by arrows in Figure 1D) to eliminate the extended Cter domain (KuΔexCter) or the entire C-terminal domain (Ku core). The biochemical properties of the purified full length Ku_{Bsub} and truncated mutants proteins were compared for interaction with $LigD_{Bsub}$ and stimulation of its activity and for interaction with various DNA substrates.

The C-terminal domain of Ku_{Bsub} is required to stimulate the ligase activity of $LigD_{Bsub}$

The primary function of Ku in bacterial NHEJ is to recruit LigD to the DNA ends thereby stimulating LigD end joining activity. We tested the stimulation of $LigD_{Bsub}$ end joining activity by Ku_{Bsub} using the full length and C-terminally truncated KuΔexCter and Ku core proteins. The Ku_{Bsub} protein variants were preincubated with a 1001 bp long linear DNA substrate carrying 5′-phosphorylated blunt ends, at a 30-fold molar excess of Ku_{Bsub} compared to DNA ends. Different amounts of $LigD_{Bsub}$ were added to the reaction mixture at a molar ratio ranging from 0.8- to 100-fold relative to DNA ends, and end joining DNA products were analyzed after 2 h incubation. Controls with Ku proteins only did not show formation of ligation products (Figure 2A), indicating that purified Ku proteins do not display any end joining activities. High concentration of $LigD_{Bsub}$ exhibited a weak end joining activity. In sharp contrast, larger amounts of multimers of the DNA substrate were formed when both Ku_{Bsub} and $LigD_{Bsub}$ proteins were present (Figure 2B), showing that Ku_{Bsub} stimulates the ligase activity of the $LigD_{Bsub}$, in agreement with a previous report (17). Increase of $LigD_{Bsub}$ concentration yielded DNA end joining products of increasing molecular weight, with a maximal ligation activity observed at 105 nM $LigD_{Bsub}$ corresponding to a ratio of 6.2 relative to DNA ends and 0.21 relative to Ku proteins. Further increase in $LigD_{Bsub}$ concentration resulted in a plateau in the ligation efficiency from 105 to 421 nM followed by a slight decrease (Figure 2E). Similarly, the KuΔexCter mutant protein stimulated $LigD_{Bsub}$ DNA end joining activity (Figure 2 panel C). However, the maximal stimulation required 8-fold higher $LigD_{Bsub}$ amount (842 nM) with KuΔexCter than with Ku_{Bsub} (Figure 2E and compare panels B and C). This result implies that the basic extended Cter region of Ku_{Bsub} plays a role in the Ku-dependent stimulation of the DNA end joining activity of $LigD_{Bsub}$. In contrast, the Ku core mutant protein was unable to stimulate $LigD_{Bsub}$ DNA end joining activity (Figure 2D).

C-terminal truncations of Ku_{Bsub} do not affect oligomerization and global architecture of the Ku core domain

To understand why Ku C-terminal truncations alter the capacity of Ku_{Bsub} to stimulate the ligase activity of $LigD_{Bsub}$, we first analyzed the structural shape of the truncated mutants and full length Ku proteins by SAXS. The different proteins were injected to a size exclusion column and eluted directly into a SAXS flow through capillary cell. SAXS experimental data were recorded from $q = 0.01$ Å^{-1} to $q = 0.4$ Å^{-1} (Figure 3A). As illustrated in Supplementary Figure S1 panel A, the Guinier extrapolations did not show aggregation phenomenons and gave unambiguous values of *Rg* for Ku_{Bsub} equal to 42.2 Å, for KuΔexCter = 34.2 Å and Ku core = 28.4 Å. The autocorrelation functions $p(r)$ were calculated from the SAXS data (Figure 3A, insert) giving information on the shape of the proteins, the maximum distance D_{max} and the *Rg* values, similar to the ones calculated from the Guinier extrapolations (Figure 3B). The $p(r)$ of the Ku core had a bell shape corresponding to a globular protein whereas the two other constructs KuΔexCter and Ku_{Bsub} presented a low decrease of the shape for the high distances. This increase of the D_{max} for the two other constructs (KuΔexCter = 150 Å and Ku_{Bsub} = 190 Å compared to Ku core = 90 Å, Figure 3B) suggested that these proteins

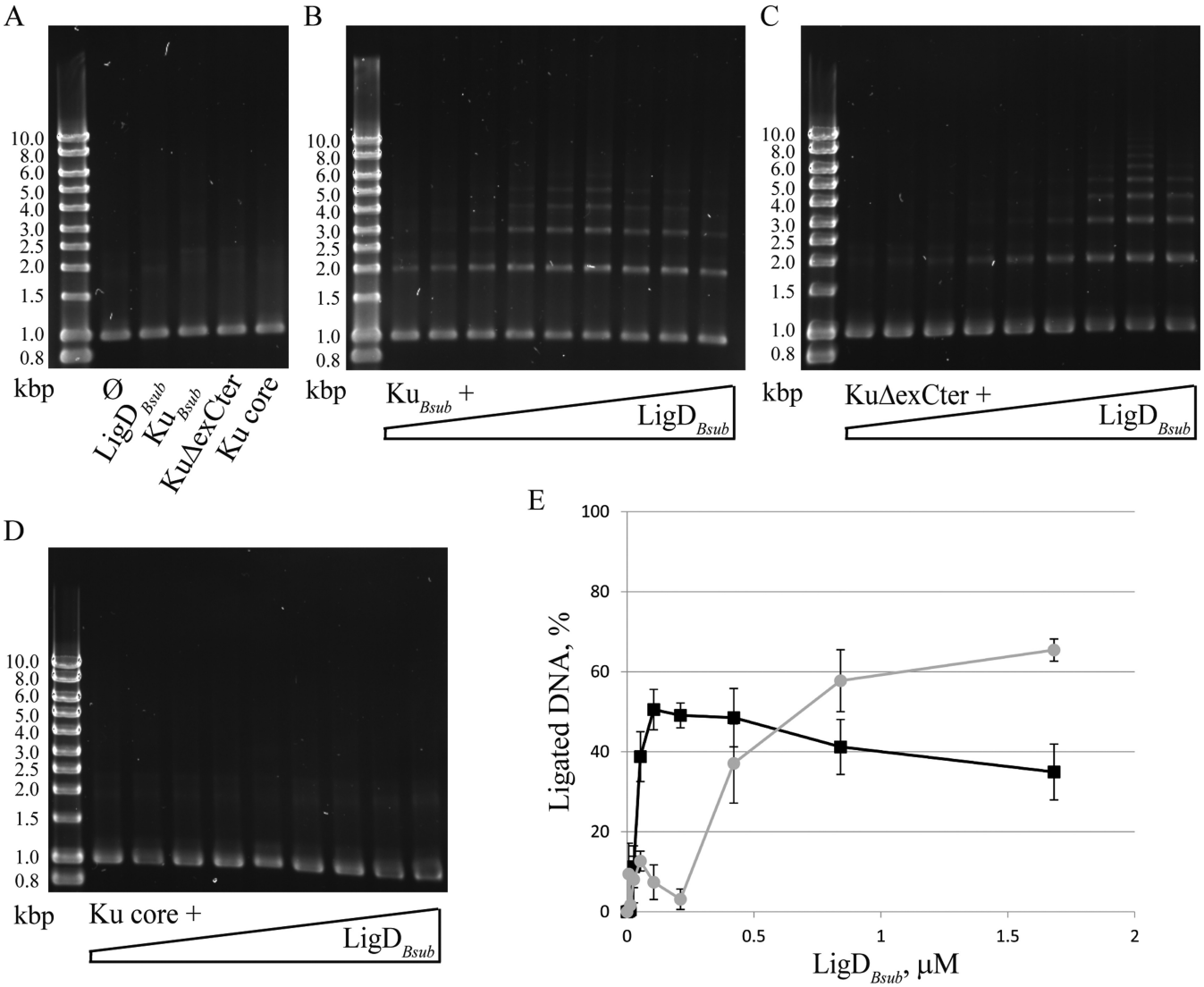
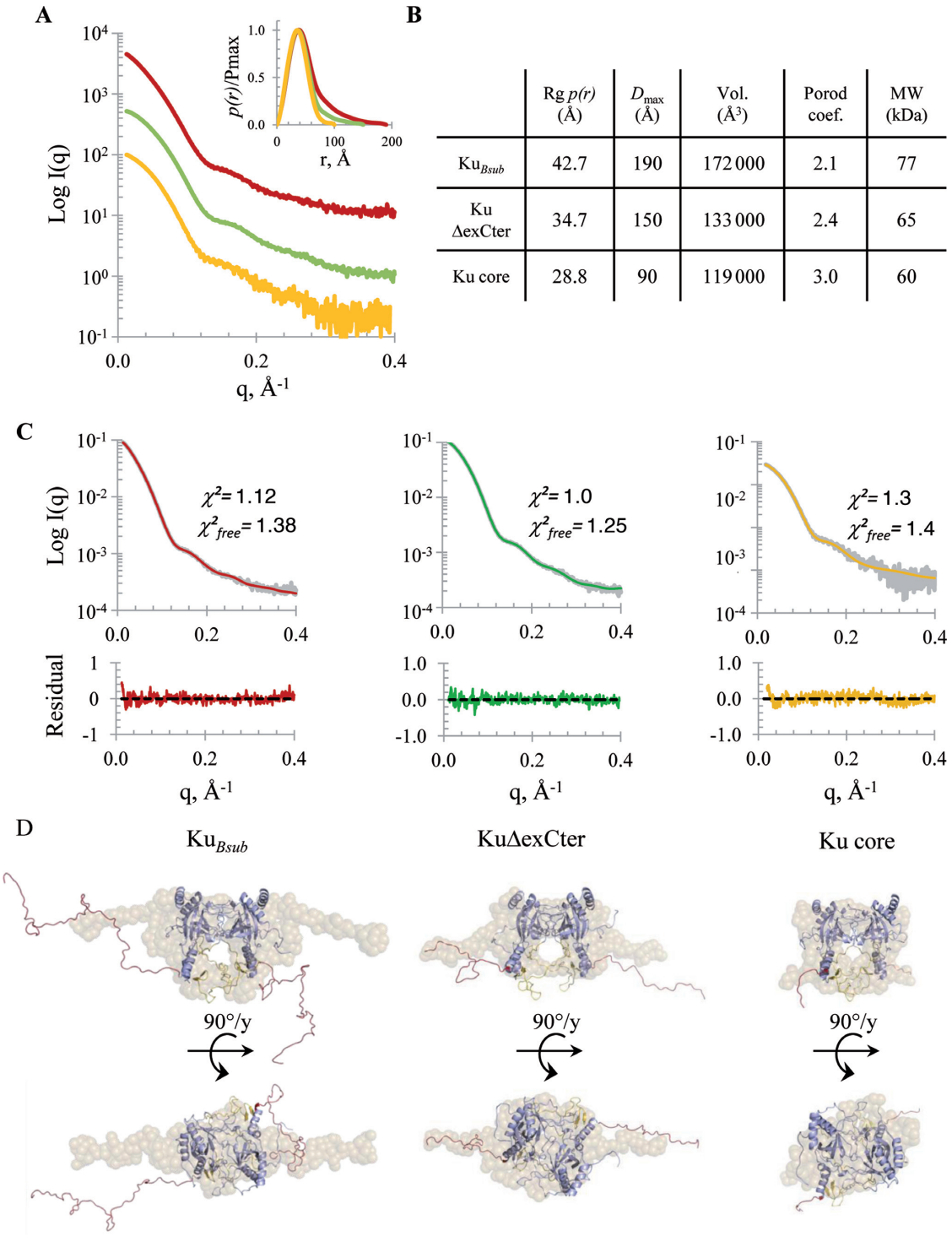


Figure 2. The C-terminal domain of Ku_{Bsub} takes part in the Ku-dependent stimulation of the ligase activity of $LigD_{Bsub}$. (**A**) Control reactions: absence of a ligase activity with Ku proteins alone. Ligation assay is described in the Material and Methods. The 5′-phosphorylated blunt-ended dsDNA product of 1001 bp (8.4 nM) is incubated alone (Ø) or with the protein indicated below the gel. Ku_{Bsub}, KuΔexCter and Ku core were at 0.5 μM, $LigD_{Bsub}$ was at 1.68 μM. Stimulation of the ligase activity of $LigD_{Bsub}$ by Ku_{Bsub}, KuΔexCter and Ku core are presented on panels **B**, **C** and **D** respectively. After preincubation of Ku_{Bsub} or truncated Ku mutants (0.5 μM) with the DNA (8.4 nM), 6.6, 13.2, 26.3, 53, 105, 211, 421, 842 or 1684 nM of $LigD_{Bsub}$ were added (from the second lane on the left of each photographed gel to the last lane on the right). Ligation products were analyzed on a SYBR Gold colored 0.8% agarose gel. The molecular weight (in kbp) of standard linear dsDNA molecules (first lane on the left) is indicated. (**E**) Quantification of the ligation efficiency of $LigD_{Bsub}$. Stimulation experiments of the ligase activity of $LigD_{Bsub}$ by Ku_{Bsub} and KuΔexCter have been repeated three times independently. Bands corresponding to the unligated DNA substrates were quantified. The percentage of DNA ligated products (calculated as (1 – the ratio of unligated molecules) × 100) was plotted against the $LigD_{Bsub}$ concentration. Error bars indicate the standard error of the mean calculated from the three independent experiments.

harbors an elongated part in the C-terminal region. The calculation of the Porod volume and the molecular weight confirmed that this increase of D_{max} was due to an elongated part and not to a change of oligomeric state. This observation was confirmed by the determination of the Porod coefficient (Figure 3B) and the normalized Kratky representation (Supplementary Figure S1B). For a globular form such as Ku core, the value of Porod coefficient is equal to 3 and the Kratky plot gives a plateau at high angles close to zero, whereas the loss of the sphericity due to the presence of the unfolded C-terminus region in Ku_{Bsub} and KuΔexCter decreases the Porod coefficient and increases the height of the plateau at high angles on the Kratky plot.

Molecular modeling from these SAXS data required to assess the oligomerization state of the different proteins. To this end, proteins were cross-linked or not with Sulfo-*EGS* and analyzed by SDS-PAGE (Supplementary Figure S4). In the absence of the cross-linker, monomeric Ku_{Bsub} (calculated molecular weight: 34 kDa), KuΔexCter (30.6 kDa) and Ku core (28.2 kDa) proteins migrated with an apparent molecular weight of ∼37 kDa, 32 kDa and 28 kDa, respectively. Addition of Sulfo-EGS led mostly to the production



	Rg $p(r)$ (Å)	D_{max} (Å)	Vol. (Å³)	Porod coef.	MW (kDa)
Ku_{Bsub}	42.7	190	172 000	2.1	77
Ku ΔexCter	34.7	150	133 000	2.4	65
Ku core	28.8	90	119 000	3.0	60

Figure 3. Biophysical parameters and low resolution structures of Ku_{Bsub} and C-terminal truncated mutants derived from SAXS. (**A**) Scattering curves for Ku_{Bsub} and for the two C-terminally truncated mutants KuΔexCter and Ku core. The experimental scattering curves in colored lines (red for Ku_{Bsub}, green for KuΔexCter and orange for Ku core) correspond to the logarithm of X-ray scattering intensity as a function of the momentum transfer $q = 4\pi\sin\theta/\lambda$, where 2θ is the scattering angle and $\lambda = 1.033$ Å is the wavelength. Autocorrelation functions $p(r)$ calculated from the SAXS data of Ku_{Bsub}, KuΔexCter and Ku core are presented in the insert. In order to compare the curves each other, the $p(r)$ are displayed as functions $p(r)/P_{max}$ versus the distance r. (**B**) Biophysical parameters calculated from experimental SAXS data of Ku_{Bsub}, KuΔexCter and Ku core. (**C**) Comparison between experimental SAXS data and all atom models of Ku_{Bsub}, KuΔexCter and Ku core. Back-calculated scattering curves from the DADIMODO modeled structures generated for Ku_{Bsub} (red line), KuΔexCter (green line) and Ku core (orange line) are fitted on the experimental scattering curves (gray lines). The calculated quality of fit, χ^2, is given for each calculated curves as well as the χ^2 free values. To illustrate the good agreement between the experimental data and the model, a residual plot is shown for the three models of Ku_{Bsub}, KuΔexCter and Ku core. (**D**) The low resolution shape generated by GASBOR (transparency sphere) and the all atom model calculated with DADIMODO (shown in cartoon) are superimposed for Ku_{Bsub}, KuΔexCter and Ku core, as indicated above models.

of proteins migrating with apparent molecular weights of 80 kDa (Ku_{Bsub}), 70 kDa (KuΔexCter) and 60 kDa (Ku core). These estimations are in good agreement with the calculated molecular weight of dimers and with biophysical parameters extracted from the SAXS data for these proteins (Figure 3B). Therefore, truncation of the C-terminal domain does not alter the dimeric state of Ku_{Bsub}.

Using scattering curves, low resolution structures of these proteins were calculated by *ab initio* modeling with GASBOR program by imposing a P2 symmetry (the choice for the P2 symmetry and comparison with low resolution structures using P1 symmetry is described in the Supplementary Results and Figure S2A). Selection of the most typical shape for each of the three Ku proteins is presented in Figure 3D (transparency spheres). The full length Ku_{Bsub} protein is composed of a central globular domain displaying a ring shape which is extended by two arms on opposite sides and excluded from the ring shape domain (Figure 3D, left panel). This ring shape domain resembles to the three dimensional (3D) structure of the human Ku70/80 heterodimer (21) where the Ku core domain of the two proteins forms a ring able to encircle a dsDNA molecule (Supplementary Figure S2B). Interestingly, truncation of the extended Cter region of Ku_{Bsub} led to a more globular protein with smaller arms and a maintained central globular domain (Figure 3D, middle panel). The Ku core protein displayed only the ring shape with a dimension similar to those of the Ku_{Bsub} and KuΔexCter proteins (Figure 3D, right panel).

Molecular modeling was performed using the known 3D structure of the human Ku protein (see Materials and Methods and Supplementary Results) and the collected SAXS data for Ku_{Bsub} and C-terminal mutants. The theoretical SAXS curves produced from these calculated atomic models were compared to the experimental SAXS data (Figure 3C). The low noisy data, the comparable values of χ^2 and χ^2 free values close to 1 and the quality of the residual plot show that the proposed molecular models are in agreement with the experimental SAXS data. Superimposition of the calculated atomic models (Figure 3D, shown in cartoons) exhibited significant agreement in size and form with the *ab initio* calculated shapes (transparency spheres). A central ring shape can be modeled for the three proteins and extended arms are visible only for Ku_{Bsub} and KuΔexCter, arms that are longer in the case of the full length Ku protein. This ring shape domain fits also well with the 3D structure of the Ku core domain of the Ku70/80 heterodimer (Supplementary Figure S3). However, during the first molecular modeling process against SAXS data of the Ku core mutant, the modification of C-terminal region by the atomic modeling algorithm DADIMODO was not sufficient to accommodate SAXS data with a good agreement (see Materials and Methods), and the same phenomenon was observed with Ku_{Bsub} and KuΔexCter. A better fit was achieved by introducing flexibility in the DNA recognition region in a second cycle of molecular modeling. DADIMODO was allowed to modify the ring structure, initially folded in beta strand in the crystallographic structure (see the model of Ku_{Bsub} in Figure 3D, shown in yellow). This modification does not affect the global structure, which always shows a ring shape, but suggests that the region of DNA recognition could be flexible in solution when no DNA is docked inside.

From these structural modeling results, we propose that the conserved ring shape of the globular domain between the three proteins corresponds to the Ku core domain of Ku_{Bsub} while the C-terminal domain is included in the arms.

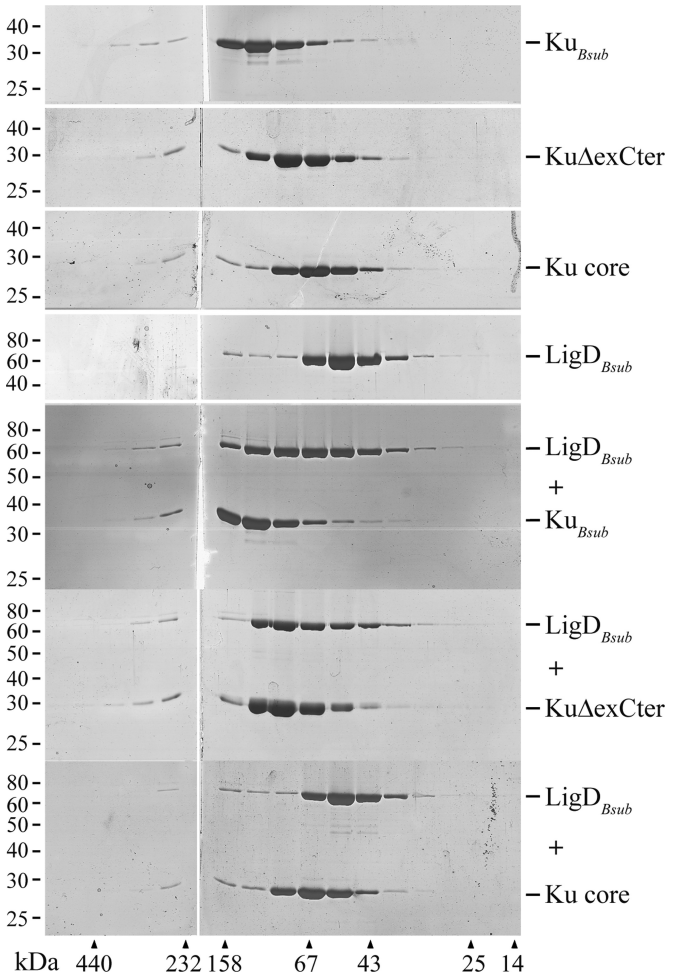


Figure 4. The minimal Cter region of Ku_{Bsub} is required for the Ku-LigD interaction *in vitro*. Ku_{Bsub} , KuΔexCter or Ku core proteins (5 nmol) and $LigD_{Bsub}$ (2.5 nmol) were injected alone or together (as indicated on the right of the photographs) onto a gel-filtration column. A total of 20 μl of eluted fractions (0.5 ml) were analyzed by SDS-PAGE. A photograph of each Coomassie blue stained gel is shown. Photographs are positioned so that eluted fractions collected at the same time after injection for each chromatography are vertically aligned. Molecular weight (in kDa) of standard proteins used to calibrate the column (arrow heads) or the gel (dashes) are indicated.

Ku/LigD interaction requires the conserved minimal Cter region of Ku_{Bsub}

LigD is thought to be recruited at DNA ends by the Ku protein which is bound to these DNA ends. Ku_{Bsub} and $LigD_{Bsub}$ are expected to interact, as Ku_{Mtub} and $LigD_{Mtub}$, which were shown to interact in a yeast two-hybrid assay (28). To evaluate the effects of the different C-terminal truncations of Ku_{Bsub} on the capacity to interact with $LigD_{Bsub}$, we performed gel filtration experiments. Determination of

elution volumes of KuΔexCter and Ku core proteins and comparison with standard proteins allowed estimation of their molecular weight: 75 kDa and 68 kDa, respectively (SDS PAGE analyzes of the eluted fraction are presented in Figure 4 and chromatograms in Supplementary Figure S5). These estimations are in good agreement with the calculated molecular weight of KuΔexCter and Ku core dimers, i.e. 61.2 kDa and 56.4 kDa, respectively. The $LigD_{Bsub}$ protein (calculated mass of 71 kDa) is eluted with an apparent molecular weight of 50 kDa, compatible with a monomeric status of the protein. In contrast, the estimated molecular weight of the dimeric Ku_{Bsub} was ∼150 kDa which is more than 2-fold higher than the dimer calculated molecular weight (68 kDa). The elongated shape of Ku_{Bsub} seen in SAXS experiment (see above) and the highly positively charged extended Cter domain could explain this apparent discrepancy.

Compared with $LigD_{Bsub}$ alone, the elution profile of the $LigD_{Bsub}$ mixed with Ku_{Bsub} or KuΔexCter proteins was more spread and was displaced toward the higher molecular masses (Figure 4 and Supplementary Figure S5), indicating that $LigD_{Bsub}$ interacts with Ku_{Bsub} and KuΔexCter. The elution volume of Ku_{Bsub} or KuΔexCter was not changed when mixed with $LigD_{Bsub}$, suggesting that Ku_{Bsub} (and KuΔexCter) forms a more globular complex when interacting with $LigD_{Bsub}$ than the protein alone. By using label-free surface plasmon resonance (SPR), the dissociation constants were determined by injecting a $LigD_{Bsub}$ solution on a sensor chip covered with one of the Ku proteins. Equilibrium dissociation constants (Kd) of 7 and 11 μM were calculated based on the binding curves of $LigD_{Bsub}$ on Ku_{Bsub} and KuΔexCter, respectively (Supplementary Figures S6A and S6B), demonstrating that the Ku_{Bsub} (and KuΔexCter)–$LigD_{Bsub}$ interaction is rather weak. The behavior of $LigD_{Bsub}$ when mixed with Ku_{Bsub} or KuΔexCter is very similar, suggesting that the extended Cter region of Ku_{Bsub} is not involved in this interaction. In sharp contrast, $LigD_{Bsub}$ did not interact detectably with the Ku core mutant as the elution profiles of the mixed Ku core and $LigD_{Bsub}$ were similar to those of the proteins alone (Figure 4 and Supplementary Figure S5). Using the SPR, in the same condition as the one used for Ku_{Bsub} and KuΔexCter, no dissociation constant could be determined (Supplementary Figure S6C). This result implies that the minimal Cter domain of Ku_{Bsub} is required for interaction with LigD. This finding explains why Ku core mutant did not stimulate the end joining activity of $LigD_{Bsub}$ (Figure 2D).

The extended C-terminal region and the core domain both contribute to Ku_{Bsub} binding to dsDNA in absence of free ends

As the decreased stimulation of $LigD_{Bsub}$ by KuΔexCter relative to Ku_{Bsub} could result from an alteration of DNA binding, we compared the DNA binding properties of Ku_{Bsub}, KuΔexCter and Ku core proteins.

These proteins were mixed with supercoiled (sc) pZE14 plasmid and binding was analyzed by EMSA as described in Materials and Methods. Two types of retardation patterns were observed, a slightly shifted and discrete band migrating just above the free scDNA substrate and a considerably shifted smear of DNA (Figure 5A). These two types of patterns were observed for Ku_{Bsub}, KuΔexCter and Ku core proteins, indicating a shared ability to bind scDNA. However, the concentration of Ku protein required to shift all DNA molecules was 0.5 μM for Ku_{Bsub} and 4 μM for KuΔexCter and Ku core, indicating that the C-terminally truncated mutants have a lower affinity for scDNA relative to the full size protein. The larger nucleoprotein complexes that migrated as smears in the gel formed with the three variants of Ku exhibited a similar difference in affinity. Removal of Ku by treatment with proteinase K (lane 10 P, Figure 5 panel A) released the DNA substrate from the different Ku/DNA complexes. The proportions of nicked circular (n) and linear (l) DNA molecules present in the pZE14 plasmid preparation remained similar after Ku binding and treatment with proteinase K, indicating that the purified Ku proteins were devoid of visible nicking or nuclease activities under these conditions.

DNA binding of Ku_{Bsub}, KuΔexCter and Ku core proteins were analyzed by TEM using 0.3 nM of supercoiled pUC19 plasmid and 30 nM of proteins (Figure 5, panels B–E). Dark field electron micrographs displayed scDNA molecules bound by each of the three proteins. In these conditions, most of the DNA molecules observed (∼90%) were bound by proteins. Two populations of complexes were observed: single DNA molecules bound by the different Ku proteins (Figure 5, panels C1, D and E), and Ku_{Bsub}–DNA networks, which were the minor part of molecules observed (10%) (Figure 5 panel C2). These networks were not observed with KuΔexCter and Ku core mutants at this concentration. These results suggest that the Ku core domain of Ku_{Bsub} can bind to DNA without free ends and that the extended Cter domain promotes DNA networks involving protein–protein or/and DNA–protein interactions. Such events are DNA dependent since Ku_{Bsub} alone (as well as Ku mutants) does not form any aggregates.

It has been proposed that the lysine rich extended Cter region of Ku_{Paer} and Ku_{Msme} could make electrostatic interactions with DNA (7,27). If this is true, adding the extended Cter to a protein without particular affinity for DNA should result in this protein gaining DNA binding ability. To assess this possibility, we fused the Ku_{Bsub}-extended Cter domain (KC) to the C-terminal domain of the Green Fluorescent Protein (GFP). The purified GFP-KC chimera and the wild-type GFP were incubated with the supercoiled pZE14 plasmid. In our conditions, this plasmid was weakly shifted by the GFP protein only at the most elevated concentration used (10 μM, Figure 5F). Remarkably, a large fraction of the scDNA was shifted by 1 μM of the GFP-KC protein (the lower concentration tested) and the DNA was entirely shifted with 3 μM. Interestingly, retardation patterns obtained with supercoiled or linear DNA by different concentrations of GFP-KC were similar with all the DNA substrate shifted at 3 μM of GFP-KC. This observation indicates that the presence of DNA ends does not change the affinity of the extended Cter region of Ku_{Bsub} for dsDNA. Altogether, these results indicate that Ku_{Bsub} binding to DNA without free ends is mediated by both the extended Cter and the Ku core domains.

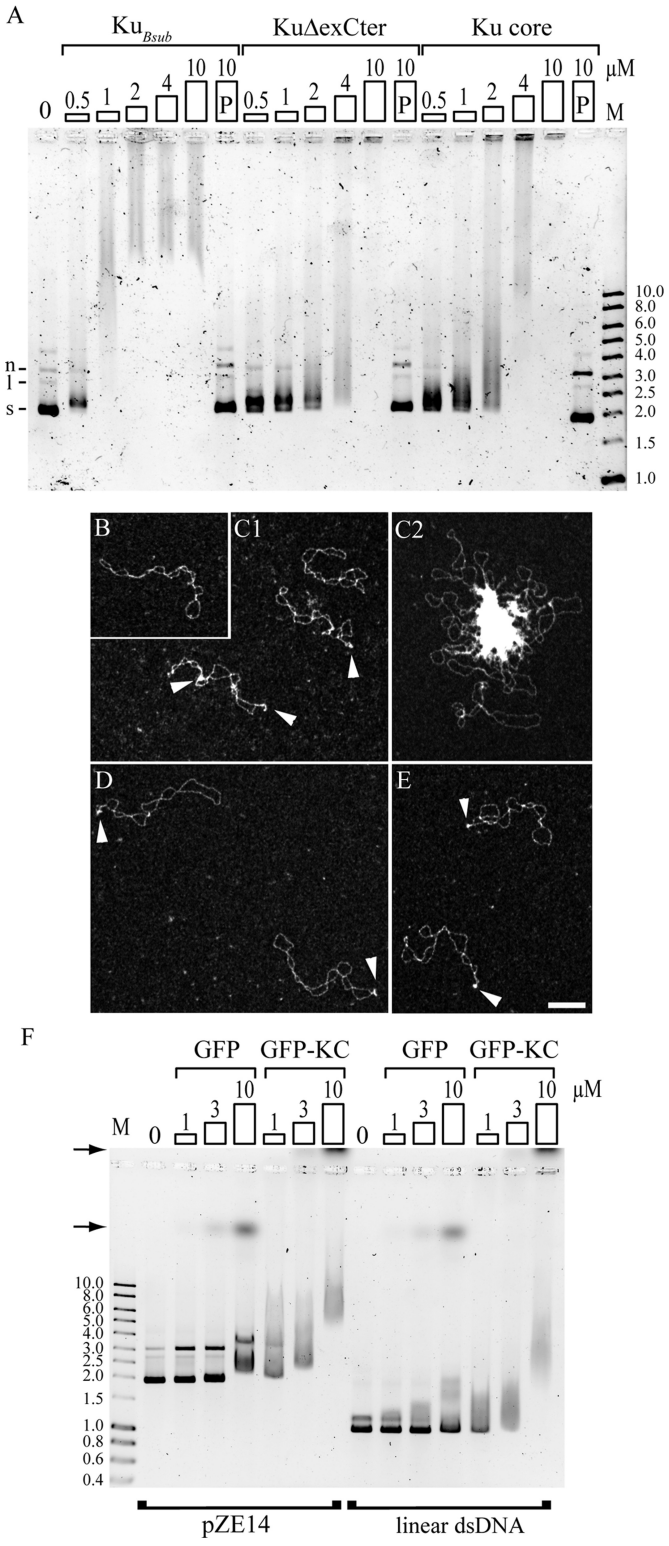


Figure 5. The extended Cter region is involved in the binding of Ku_{Bsub} on supercoiled DNA and gives the DNA binding ability to the GFP. (**A**) Binding of Ku_{Bsub}, KuΔexCter or Ku core proteins on the supercoiled pZE14 plasmid analyzed by EMSA. Increasing amounts of Ku_{Bsub}, KuΔexCter or Ku core proteins (concentrations (in μM) and proteins are indicated above photographs of the gels) were incubated with 50 ng (1.4 nM) of the pZE14 plasmid and loaded onto a 0.7% agarose gel. Photographs of the gels stained by SYBR Gold after electrophoresis are shown. The migration pattern of the DNA molecules without proteins is shown in the lane 0, where the main band corresponds to the supercoiled pZE14 plasmid (s) and faint bands, above the supercoiled plasmid, to the nicked (n) and linear (l) forms of this plasmid. To exclude a putative nuclease contamination in the purified proteins, a second reaction mixture containing 10 μM of Ku_{Bsub} or one of the truncated protein and the DNA was prepared. After incubation, proteinase K (0.5 mg/ml) and SDS (0.2%) were added and the mixture was incubated at 50°C for 30 min prior to be loaded onto the gel (indicated by a 'P' in a box above the gel). M: dsDNA molecular weight marker with sizes indicated on the side (in kbp). (**B–E**) Electron microscopy analysis of Ku_{Bsub} and Ku C-terminal mutants in complex with supercoiled dsDNA. Dark field electron micrographs of 0.3 nM of supercoiled pUC19 plasmid alone (**B**) or in presence of 30 nM of Ku_{Bsub} (C1–2), KuΔexCter (D) or Ku core (E) proteins in binding buffer (10 mM Tris-HCl, pH 8, 50 mM NaCl). Ku_{Bsub} and C-terminal truncated proteins bind to supercoiled DNA (white arrows) without DNA covering. The protein–protein and/or protein–DNA interactions induce the formation of nucleoprotein networks with Ku_{Bsub} (C2), which are not observed with both Ku mutants (**D** and **E**). Bar represents 100 nm. (**F**) The GFP fused to the extended Cter region of Ku_{Bsub} (GFP-KC) binds supercoiled and linear dsDNA. Increasing concentrations of GFP or GFP-KC (as indicated above the gel) were mixed with 50 ng of the supercoiled pZE14 (1.4 nM) or the 1001 bp linear dsDNA with 5′-phosphorylated blunt ends (3.8 nM), as indicated below the gel. Analysis of nucleoprotein complexes formed was done as described in panel A. Bands pointed by an arrow correspond to the fluorescent GFP (lower band) and GFP-KC proteins (upper band).

Ku_{Bsub} and C-terminally truncated Ku proteins thread inward linear DNA molecule after entering via the ends

We tested the ability of Ku_{Bsub} and truncated mutants to bind linear dsDNA. A blunt-ended 5′-phosphorylated 279 bp DNA substrate (a DNA molecule without 5′-phosphorylated ends gave identical results, not shown) was incubated with each of the three Ku protein variants. Ku_{Bsub} and both truncated mutants led to the formation of multiple discrete retarded DNA bands showing that all these proteins were able to bind linear dsDNA molecules (Figure 6A). The molecular weight of Ku–DNA complexes increased with increasing protein concentration. At 25 and 50 nM of proteins the number of retarded species was 3 and 4, respectively for all protein variants, suggesting that they display similar affinity for DNA ends. Increasing the concentration up to 400 nM yielded a single thick shifted band for Ku core and KuΔexCter and nucleoprotein complexes that did not enter the gel for Ku_{Bsub}. This result shows that DNA molecules with free ends can bind several C-terminally truncated Ku proteins until a maximum is reached. Proportions of unbound DNA were 60, 58 and 40% with 50 nM of Ku_{Bsub}, KuΔexCter and Ku core proteins, respectively. Interestingly, when the concentration of protein was doubled, the fraction of unbound DNA remained similar whereas the number of retarded species increased and the amount of DNA in the two less retarded bands clearly decreased on the gels. This observation suggests that the DNA binding of Ku_{Bsub} and both truncated mutants at DNA ends is cooperative, as previously described for the human Ku (49).

To investigate the nature of these nucleoprotein complexes, we analyzed the binding of the three Ku protein variants on a linear dsDNA by TEM (Figure 6, panels B–J). Incubation of 30 nM of Ku_{Bsub} with 0.3 nM of a 1876 bp linear substrate promoted the formation of two protein–DNA complex populations: (i) single DNA molecules with several Ku_{Bsub} proteins predominantly covering DNA ends

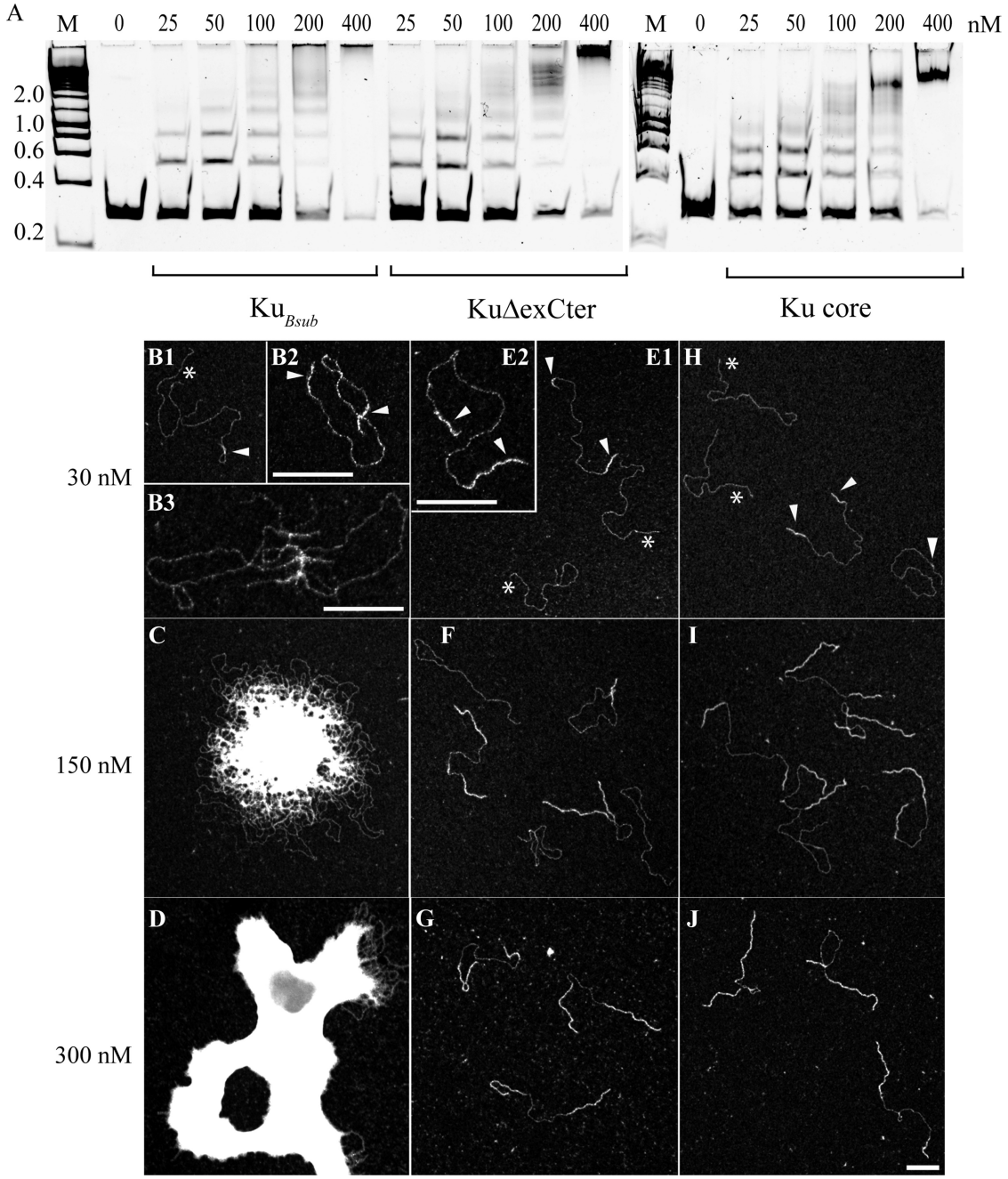


Figure 6. Ku$_{Bsub}$ and C-terminal truncated Ku mutants interact with linear dsDNA. (**A**) Binding of the different Ku proteins on a 279 bp linear dsDNA with 5′-phosphorylated blunt ends. The experiment was done as described in the legend of Figure 5 panel A except that the nucleoprotein complexes formed with 50 ng (13.7 nM) of the 279 bp substrate were analyzed on a 5% polyacrylamide gel. Note that concentrations are in nM. M: dsDNA molecular weight marker with sizes indicated on the side (in kbp). (**B–J**) Electron microscopy analysis of Ku$_{Bsub}$ and Ku C-terminal mutants in complex with linear dsDNA. (**B–D**) Dark field electron micrographs of Ku$_{Bsub}$, (**E**–**G**) KuΔexCter and (**H–J**) Ku core proteins at different concentrations (B, E, H: 30 nM, C, F, I: 150 nM and D, G, J: 300 nM) in presence of 0.3 nM of a 1876 bp linear dsDNA molecule with 5′-phosphorylated blunt ends in binding buffer (10 mM Tris-HCl, pH 8, 50 mM NaCl). In micrographs obtained with 30 nM of proteins, three kinds of molecules were observed: naked DNA molecules (asterisks), Ku–DNA complexes (pointed by white arrows) formed mostly at DNA ends and nucleoprotein networks caused by protein–protein and/or by protein–DNA interactions (**B3**). Size of these networks increased with Ku$_{Bsub}$ concentration for finally form characteristic structures of DNA condensation (C and D). These nucleoprotein networks were never observed with the truncated mutants (E–J) whereas the number of Ku–DNA complexes increases (all DNA molecules are bound at 300 nM) and the covering of DNA molecules progressed. Bars represent 100 nm.

(white arrows on Figure 6B1 and B2), demonstrating that Ku_{Bsub} is loaded onto DNA ends, and (ii) Ku_{Bsub}–DNA networks mediated by intra- and intermolecular bridging events based on protein–protein or/and DNA–protein interactions (Figure 6B3, Supplementary Figure S7). These bridging events led to networks formation and concentration of DNA ends at the heart of these networks (Figure 6C and Supplementary Figure S7). An increase of Ku_{Bsub} concentration increased the size of these networks and decreased the number of Ku_{Bsub}–DNA molecules not trapped into networks. At 300 nM, the networks led to the formation of aggregates (Figure 6D). For both truncated Ku proteins, we did not observe protein–DNA networks but only DNA molecules covered by Ku mutants, predominantly at the ends, and naked DNA molecules (Figure 6, panels E1–J). This suggests that the extended Cter domain mediates the assembly of Ku_{Bsub}–DNA networks. Remarkably, the length of KuΔexCter threads increased with protein concentration until the threads covered nearly all the DNA molecule (compare panels E1 with G). The Ku core protein displayed the same behavior as KuΔexCter (Figure 6, panels H–J). These results demonstrate that the C-terminally truncated Ku proteins are loaded onto DNA ends, as observed with full size Ku_{Bsub}, and they reveal inward threading of Ku onto linear DNA molecules, which could not be easily observed with full size Ku_{Bsub} due to the formation of Ku_{Bsub}–DNA networks at elevated concentrations.

In light of these results, the multiband pattern observed in EMSA experiments (Figure 6A) could be interpreted as progressive inward threading of Ku protein variants onto the DNA with Ku entering via the DNA ends. The single retarded band obtained with 400 nM of the C-terminally truncated proteins could correspond to DNA molecules completely covered. The absence of visible retarded bands at 400 nM of Ku_{Bsub} could be due to the formation of high molecular weight nucleoprotein complexes or aggregates, as observed in Figure 6 (panels C and D).

Truncation of the extended C-terminal region alters Ku_{Bsub} ability to bridge linear DNA molecules

The ability of Ku_{Bsub} to interact and form Ku_{Bsub}–DNA networks with supercoiled and linear DNA molecules prompted us to test whether Ku_{Bsub} was able to bridge two DNA molecules, a potentially important property to maintain both ends of a DNA break in close proximity. Using a previously described bridging assay (46), we generated a blunt-ended 1001 bp linear DNA molecule 5′-biotinylated at one end and 5′-phosphorylated at the other end (1001BP in Figure 7) that was attached to streptavidin coated magnetic beads. Ku_{Bsub} was then loaded on the DNA substrate and the ability of the Ku_{Bsub}–1001BP complex to form a bridge with another DNA molecule (5′-phosphorylated blunt-ended 1876 bp DNA fragment) was assessed by analyzing the DNA recovered from the beads on agarose gel as described in Materials and Methods (see Figure 7 panel B). Non-specific binding of DNA on streptavidin-coated beads was measured using a DNA fragment carrying 5′-phosphorylated ends on both extremities (1001PP, Figure 7B). After removing unbound molecules, we also analysed the protein content of beads after the incubation of Ku_{Bsub} with the 1001BP substrate ('M' fraction in Figure 7A) and after the incubation of the 1876 bp ('E' fraction in Figure 7A).

With Ku_{Bsub}, a larger amount of the 1876 bp DNA fragment was retained on beads attached to 1001BP compared to 1001PP (Figure 7B), while a fraction of the protein was associated only with beads that were attached to the 1001BP substrate (Figure 7A). This indicates that Ku_{Bsub} can bridge DNA molecules. In contrast, the KuΔexCter and Ku core mutant proteins did not exhibit any detectable bridging activity as similar amounts of the 1876 bp DNA fragment were retained on beads attached to 1001BP or 1001PP (Figure 7B), under conditions where both mutant proteins associated only with beads attached to 1001BP (Figure 7A). These results are consistent with the TEM results and confirm that the extended Cter of Ku_{Bsub} is required to maintain DNA extremities in close proximity.

The extended C-terminal region acts to limit threading and to concentrate Ku_{Bsub} at DNA ends

The TEM analysis of micrographs obtained with 30 nM of Ku proteins on linear dsDNA suggested that the nucleofilaments formed by KuΔexCter and Ku core were longer than those observed by Ku_{Bsub} (compare panels B, E and H in Figure 6). However, the formation of Ku–DNA networks even at this Ku_{Bsub} concentration did not permit to quantify this difference. We searched for reaction conditions limiting the formation of Ku_{Bsub}–DNA networks while maintaining threading inward the DNA molecules via their ends.

The increase of ionic strength (from NaCl 50 mM to 200 mM) resulted in a significant decrease of Ku_{Bsub}–DNA networks (Supplementary Figure S8). In this reaction condition and at a concentration of 150 nM of Ku protein variants, we monitored the formation of Ku_{Bsub}–DNA networks and aggregates (such as those presented in panels B3, C and D of the Figure 6). Because the number of molecules contained in these high molecular weight complexes cannot be defined, their kinetic of formation was measured by counting the number of naked DNA molecules and nucleoprotein complexes not trapped into these networks over time (called untrapped DNA molecules, untrap. DNA in Figure 8D). We found that the frequency of untrapped DNA molecules with Ku_{Bsub} decreased significantly (Mann–Whitney U test, $P < 0.05$) by up to 50% of the initial quantity after 60 min of incubation, whereas this frequency remained stable after incubation with KuΔexCter and Ku core for the same time. For incubation times longer than 15 min, the frequency of untrapped DNA molecules was significantly lower with Ku_{Bsub} than with KuΔexCter and Ku core (Mann–Whitney U test, $P < 0.05$). Thus, the formation of high molecular weight nucleoprotein networks mediated by Ku_{Bsub} depends on the concentrations of salt and protein as well as on the incubation time.

We classified untrapped DNA molecules into four types: (i) naked DNA, (ii) complexes with proteins at one or (iii) both DNA ends, and (iv) complexes with protein inside the DNA molecule with one, two or without covered ends (Figure 8E). After 2 min of incubation, the majority of DNA molecules (>83%) were bound by Ku_{Bsub}, KuΔexCter and Ku core, at one or mainly both ends. A small percentage

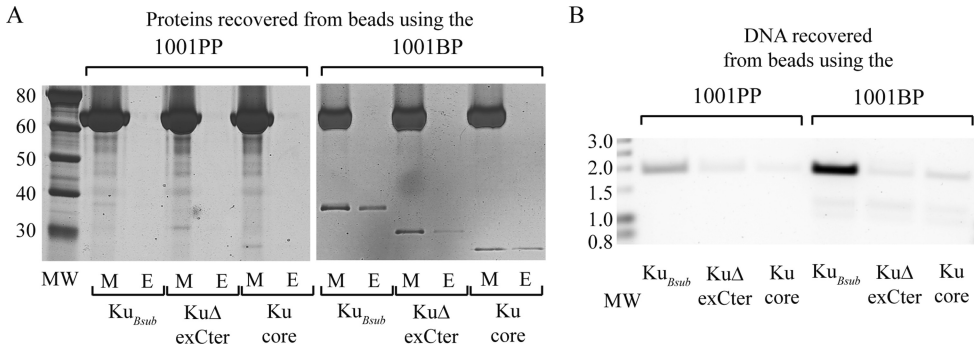


Figure 7. An efficient bridging of linear DNA molecules requires the extended Cter of Ku_{Bsub}. The DNA bridging assay was performed as described in the Material and Methods. (**A**) Protein contents of streptavidin coated magnetic beads were analyzed by SDS PAGE from samples taken: M, after Ku_{Bsub} (34 kDa), KuΔexCter (30.6 kDa) or Ku core (28.2 kDa; as indicated below photographs of the Coomassie blue stained gel) loading on the 1001PP (control experiment) or the 1001BP dsDNA fragment, and E, after incubation of the 1876 bp DNA molecule and washes of beads. The protein migrating between the 60 and 80 kDa size-markers corresponds to the BSA. (**B**) DNA recovered from beads incubated with one of the 1001 base pairs substrate, one of the Ku proteins (as indicated above and below the SYBR Gold stained agarose gel) and the 1876 bp DNA fragment. MW: molecular weight markers with sizes indicated on the side (kDa in panel A and kbp in B).

(3%) of the DNA molecules displayed proteins bound inside (see white arrow in Figure 8A, termed diffusion events in Figure 8E). In at least two out of three cases, these diffusion events were concomitant with one or both DNA ends covered, suggesting that threads of Ku were broken during sample preparation and/or that part of the thread diffused inside the DNA molecule. Interestingly, we observed that the population of DNA molecules bound at one or two ends by Ku_{Bsub} was stable for 30 min (82 ± 7%) and decreased. The observation that the frequency of untrapped DNA molecules with Ku_{Bsub} also decreased over time suggests that the DNA molecules covered by Ku_{Bsub} at one or both DNA ends are preferentially trapped into Ku_{Bsub}–DNA networks.

We determined the fraction of DNA covered by the three Ku proteins in untrapped DNA molecules (ratio of the sum of DNA length covered by Ku_{Bsub} or truncated proteins over the total length of DNA molecule at each incubation time). The distribution is shown on Figure 8F. The mean fractions of DNA covered were 20 ± 6; 37 ± 13; 38 ±13; 48 ±13% for KuΔexCter and 18 ± 5; 34 ± 6; 34 ± 7; 45 ± 7% for Ku core, for 2, 15, 30 and 60 min of incubation time, respectively. Interestingly, these values were significantly lower with the Ku_{Bsub} protein (Mann–Whitney U test, $P < 0.05$) and demonstrated a more limited threading (13 ± 7; 17 ± 7; 23 ± 7; 20 ± 8%). In addition, for incubation times longer than 15 min, more than 50% of bound DNA molecules displayed ends covered on a length ranging from 250 bp up to 570 bp for both truncated proteins, whereas this proportion fell to 13% for Ku_{Bsub} with a covered length ranging from 250 bp to a maximum of 480 bp. These results indicate that the extended Cter domain limits inward DNA threading of Ku_{Bsub}.

The properties to bind to DNA ends and to thread inward the DNA molecule should allow Ku_{Bsub} and C-terminally truncated Ku proteins to protect DNA from degradation by an exonuclease, as previously described for Ku from different bacterial species, but also to protect against cleavage by an endonuclease. Moreover, the limited threading capacity of full length Ku_{Bsub} relative to that of KuΔexCter, as observed by electron microscopy, could lead to different DNA protection efficiencies. We probed the capacity of Ku_{Bsub} and KuΔexCter proteins to protect DNA molecules from degradation by the T5 exonuclease. The 5′-phosphorylated blunt-ended 1001 bp linear dsDNA substrate was preincubated with increasing amounts of Ku_{Bsub} or KuΔexCter proteins and then subjected to the T5 exonuclease. Analyses of the DNA degradation patterns showed that Ku_{Bsub} and KuΔexCter were equally efficient to inhibit the T5 exonuclease and protect the DNA ends (Figure 9A). This 1001 bp DNA molecule harbors two identical *HincII* sites, one positioned at 266 bp (hereafter named site 266) from one end and the other (named site 421) at 421 bp from the same DNA end (Figure 9B). DNA substrate pre-bound with Ku_{Bsub} or KuΔexCter was digested by HincII and analyzed on an agarose gel (Figure 9C). In the control reaction mixture where Ku proteins were omitted, the three main DNA bands expected from complete digestion of the two HincII sites were obtained (lane 2 in Figure 9C). Prebinding of Ku_{Bsub} or KuΔexCter on the DNA molecules led to several results. Firstly, Ku_{Bsub} and the truncated mutant protein protected the DNA from HincII cleavage, as indicated by the absence of digested DNA molecules at elevated protein concentrations (Figure 9C, lanes 8, 13, 14). Secondly, the concentration of KuΔexCter required to protect 50% of the linear DNA substrate is approximately two times lower compared to Ku_{Bsub} (around 1 μM versus 2 μM, see quantification in Supplementary Figure S9B). Thirdly, quantification of DNA amounts in bands corresponding to the 580 bp and the 266 bp HincII digested products provided a good estimation of the number of cutting events at site 421 and at site 266, respectively (Figure 9D). The number of cutting events decreased more rapidly at the site 266 than at the site 421 when raising the concentration of Ku_{Bsub} as well as of KuΔexCter (compare dashed lines to full lines in Figure 9D). This result suggests that the site 266, the site closest from the end, was preferentially protected by Ku_{Bsub} (and, to a lesser extent, by KuΔexCter mutant), most probably by its threading from this end. Although the site 266 seemed to be protected with the same efficiency by Ku_{Bsub} and the C-terminal mutant, this is not the case for the site 421. Indeed, lower concentrations of KuΔexCter are

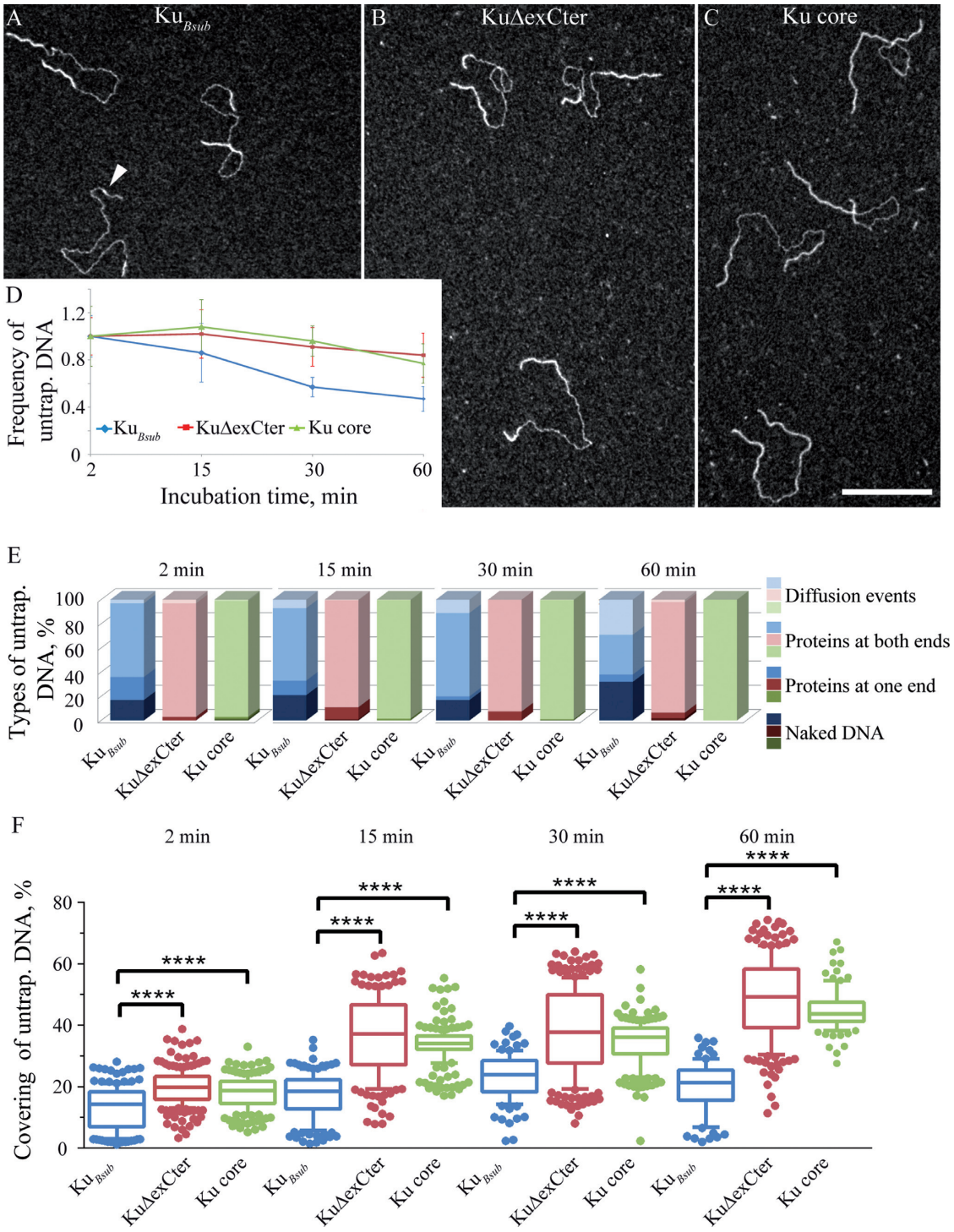


Figure 8. C-terminal truncated Ku mutants thread more efficiently inward the linear dsDNA molecule. Dark field electron micrographs of nucleoprotein assemblies formed by a 1876 bp linear dsDNA molecule with 5′-phosphorylated blunt-ends and 150 nM of (**A**) Ku_{Bsub}, (**B**) KuΔexCter or (**C**) Ku core proteins in the binding buffer with an increased salt concentration (200 mM NaCl). Bar represents 200 nm. (**D**) Linear representation of the frequency of untrapped DNA molecules (corresponding to naked DNA and DNA covered by proteins but not trapped into networks, untrap.). After 2 min of incubation, no DNA molecules trapped into networks or aggregates was observed with the three Ku proteins. The frequency of untrapped DNA molecules was calculated by dividing the number of untrapped DNA molecules per μm^2 at the indicated time of incubation by the number of these molecules per μm^2 at 2 min. Two hundred untrapped DNA molecules were analyzed for each incubation time. Error bars indicate the standard deviation of the mean. (**E**) Histogram representation of types of untrapped DNA molecules for each Ku proteins: naked DNA molecules, complexes with proteins at one or both ends and complexes with proteins inside the DNA molecule and one, two or without covered ends (named diffusion events, an example of which is pointed by a white arrow in panel A). Two hundred untrapped DNA molecules were observed at different incubation times (indicated above the histograms) with the different Ku proteins (indicated on the abscissa). (**F**) Statistical analysis of the DNA length covered by Ku_{Bsub} and both truncated mutants over the time of incubation. Two hundred untrapped DNA molecules were analyzed. ****: significant difference with Mann–Whitney U test, P-value < 0.0001.

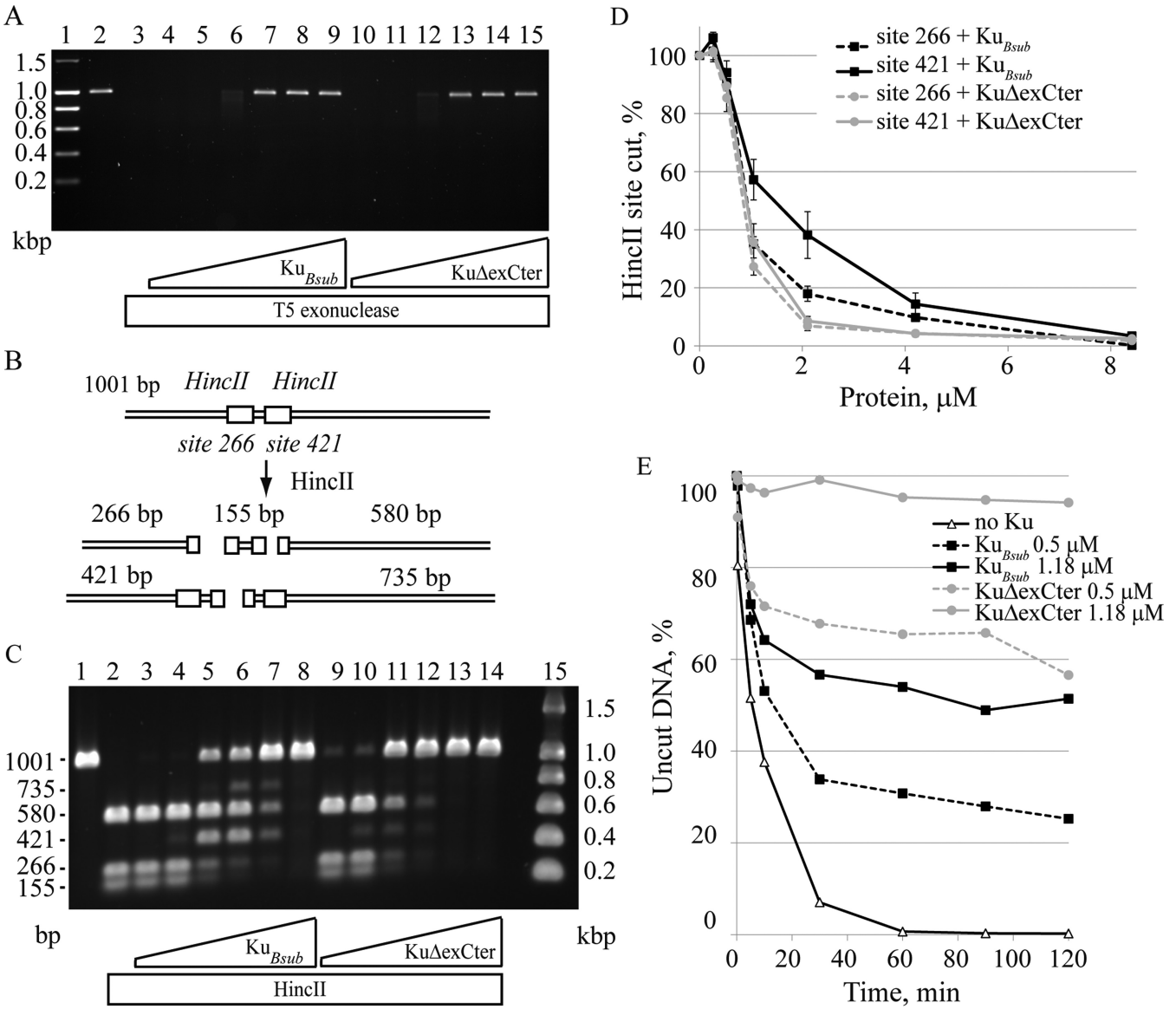


Figure 9. Ku_{Bsub} and KuΔexCter proteins protect a linear dsDNA from exo- and endonuclease activities. (**A**) Protection of DNA ends by Ku_{Bsub} and KuΔexCter proteins against the T5 exonuclease. A 1001 bp linear dsDNA with 5′-phosphorylated ends (8.4 nM) was incubated in buffer L, 30 min at 37°C without (lanes 2 and 3) or with increasing amounts of Ku_{Bsub} (0.13, 0.26, 0.51, 1.02, 2.04 and 4.08 μM corresponding to lane 4 to lane 9) or KuΔexCter (same concentration as Ku_{Bsub}; lane 10 to lane 15) proteins. T5 exonuclease (0.5 U) was then added in each reaction mixtures (lane 3 to lane 15) for 30 min at 37°C. Digested products were analyzed as described in the Materials and Methods. Lane 3: control experiment showing the complete degradation of the DNA molecule by the T5 exonuclease without other proteins added. (**B**) Schematic representation of the 1001 bp linear dsDNA with 5′-phosphorylated blunt ends and localization of the two *HincII* sites (rectangle) present on this molecule. All possible digested molecules obtained after a complete or a partial HincII digestion are shown. (**C**) Protection of internal DNA sites by Ku_{Bsub} and KuΔexCter. Preincubation of the 1001 bp DNA molecule was done as described in A, without (lane 1 and 2) or with increasing amounts of Ku_{Bsub} (0.26, 0.52, 1.05, 2.1, 4.2 and 8.4 μM, lane 3 to lane 8) or KuΔexCter (same concentrations as Ku_{Bsub}; lane 9 to lane 14). HincII restriction enzyme (4.8 U) and BSA (0.1 mg/mL) were added (lanes 2 to 14) or not (lane 1) to the reaction mixtures for 2 h at 37°C. Lane 1 panel A and Lane 15 panel C: dsDNA molecular weight marker with sizes indicated on the side (in kbp). (**D**) Quantification of the Ku mediated protection at each *HincII* sites on the 1001 bp dsDNA substrate. Amount of DNA in bands corresponding to the 266 and 580 bp digested products was quantified in three independent experiments performed as in panel C. The percentage of cuts at site 266 (dashed lines) and at site 421 (full lines) was obtained by dividing the number of the 266 and 580 bp digested molecules, for each protein concentration tested (Ku_{Bsub}, black lines or KuΔexCter, gray lines), by the corresponding number of molecules obtained in the reaction mixture containing only the HincII protein (panel C lane 2) and multiplying by 100. Error bars indicate the standard error of the mean calculated from the three independent experiments. (**E**) Kinetics of HincII digestion of the 1001 bp dsDNA substrate preincubated with different Ku concentrations. Preincubation of the 1001 bp DNA molecule was done as described in A, without (no Ku) or with 0.5 or 1.18 μM of Ku_{Bsub} or KuΔexCter as indicated. HincII enzyme (3.84 U) and BSA (0.1 mg/ml) were added and incubated for 2 h at 37°C. Samples of the reaction were extracted and the reaction was stopped at 20", 5', 10', 30', 60', 90' and 120' after addition of HincII. DNA molecules were analyzed as described in the Materials and Methods. The percentage of uncut DNA was determined by dividing the quantity of the 1001 bp DNA by the total amount of DNA in each samples and multiplying by 100. This percentage of uncut DNA is plotted against the time of the sample extraction.

required to protect this site compared to Ku_{Bsub} (compare grey full line to black full line in Figure 9D). In keeping with the electron microscopy experiments, this difference is consistent with the more limited threading capacity of Ku_{Bsub} relative to the KuΔexCter mutant.

The kinetic of DNA protection against HincII cleavage was determined under conditions where Ku_{Bsub} and KuΔexCter concentrations prevent total digestion after 2 h of incubation. Reaction samples were extracted at different times and analyzed to determine the fraction of uncut 1001 bp DNA molecule. In our testing conditions, HincII digested nearly all the 1001 bp DNA molecules in 30 min (Figure 9E). When 0.5 μM of Ku_{Bsub} was pre-incubated with the DNA, two phases was observed during the digestion kinetic. The first one is a rapid digestion of the DNA substrate in the first 30 min after addition of the HincII enzyme, leading to 34% of uncut DNA. Beyond 30 min, a very slow digestion phase was observed, leading to 25% of undigested DNA after 120 min. The same biphasic digestion kinetic was observed using 1.18 μM of Ku_{Bsub} except that the proportions of uncut DNA were 57% and 51% after 30 and 120 min, respectively (Figure 9E). While 1.18 μM of KuΔexCter totally protected the DNA, a quite stable DNA protection was also measured with 0.5 μM of KuΔexCter beyond 30 min, after a first phase of rapid digestion leading to 68% of uncut DNA (Figure 9E). Based on the observations presented above, these results strongly suggest that during the first phase of the kinetic, DNA molecules displaying *HincII* sites uncovered by nucleoprotein filaments formed after the pre-incubation with Ku_{Bsub} (or KuΔexCter) were rapidly cut by the endonuclease while those bearing *HincII* sites covered by Ku_{Bsub} (or KuΔexCter) remained mostly protected over the time. Then, the proportion of uncut DNA molecules observed during the second phase should reflect the proportion of DNA molecules with *HincII* sites protected by nucleoprotein filaments formed during pre-incubation with Ku_{Bsub} or KuΔexCter. Consequently, nucleoprotein filaments formed by Ku_{Bsub} and KuΔexCter appeared equally stable. Moreover, the higher amount of DNA molecules fully protected by 0.5 μM of KuΔexCter compared to 0.5 μM of Ku_{Bsub} indicates that nucleoprotein filaments formed by KuΔexCter were longer than those formed by Ku_{Bsub} at the same concentration, reinforcing results obtained by TEM experiments (Figure 8F).

Ku_{Bsub} and KuΔexCter were also able to protect supercoiled dsDNA (not shown) and nicked circular dsDNA against cleavage by HincII at elevated concentrations (Supplementary Figure S9 panel A). However, for both proteins protection of the 1001 bp linear DNA molecule was more efficient than protection of the nicked circular molecule (Supplementary Figure S9 panel B). Moreover, the concentrations of KuΔexCter and Ku_{Bsub} needed to protect 50% of the circular DNA are quite similar (~3 μM and 3.5 μM, respectively, Supplementary Figure S9B). These results support the notion that the higher protection of linear DNA by KuΔexCter relative to Ku_{Bsub} is due to a more efficient threading from DNA ends. Protection of circular DNA molecules by Ku_{Bsub} and KuΔexCter could be related to the formation of large nucleoprotein complexes at elevated concentrations (see Figure 5A) which could block accessibility of HincII sites.

DISCUSSION

While the lysine rich extended C-terminal regions of Ku_{Paer} and Ku_{Msme} are involved in the interaction of Ku with DNA, the role of this interaction in the NHEJ pathway remains unclear (7,27). In this study, we report that the C-terminal domain of Ku_{Bsub}, which was delineated by an *in silico* analysis of prokaryotic Ku proteins, is composed of two regions (Figure 1) with distinct functions. The minimal C-terminal domain following the Ku core domain is required for interaction with $LigD_{Bsub}$ (Figure 4) and for the Ku-dependent stimulation of the ligase activity (Figure 2). Next, the lysine rich extended C-terminal region increases the Ku-dependent stimulation of $LigD_{Bsub}$ (Figure 2). Characterization of the DNA binding properties of Ku_{Bsub} truncated mutants shows that the extended Cter region is required for the formation of Ku_{Bsub}–DNA networks with linear as well as circular dsDNA (Figures 5 and 6) and for the bridging of DNA molecules (Figure 7). Based on the SAXS analysis and molecular modeling of the dimeric Ku_{Bsub}, showing a ring shape with extended C-terminal arms, electron microscopy analyses of the Ku_{Bsub}–DNA interaction and DNA protection against endonuclease cleavage (Figures 3, 6 and 9), we propose that Ku_{Bsub} binds dsDNA ends and threads inward the DNA molecule, as its eukaryotic homologue. Finally, we demonstrate that the basic extended Cter domain acts to limit threading of Ku_{Bsub} inward the dsDNA (Figures 8 and 9).

Taken together, these results lead us to propose several roles for the Ku C-terminal domain of *B. subtilis*. Firstly, its affinity for linear as well as circular DNA allows the recruitment of Ku_{Bsub} to unbroken DNA, placing Ku_{Bsub} in a suitable position for efficiently handling DNA ends in the cell, as already suggested for the Ku_{Msme} protein (27). When a DSB occurs, Ku binds the broken ends and threads inward the chromosome, a process requiring the addition of new Ku_{Bsub} molecules at the DNA ends. Secondly, the extended Cter domain limits the extent of threading probably by interacting with the DNA onto which Ku_{Bsub} is loaded (see Supplementary Figure S7 panel E; other explanations for inhibition of Ku threading are discussed below). Finally, $LigD_{Bsub}$ will be recruited at the DSB site by the minimal Cter region of Ku_{Bsub} proteins bound near DNA ends. Thus, by limiting threading, the extended Cter domain favors accumulation of Ku_{Bsub} within a short distance of the DSB and the recruitment of $LigD_{Bsub}$ at this site for repair.

Based on our data, other roles could be proposed for the extended Cter domain of Ku_{Bsub}. The formation of Ku_{Bsub}–DNA networks (see electron micrographs Figure 6 panels B3 and C) suggests that the extended Cter could also play an active role in bringing DNA molecules together, potentially maintaining DSBs in close proximity to facilitate sealing by $LigD_{Bsub}$. This hypothesis is reinforced by the requirement of the extended Cter to bridge two dsDNA molecules in absence of $LigD_{Bsub}$ (see Figure 7). However, the fact that KuΔexCter is only slightly altered in its ability to stimulate the ligase activity of $LigD_{Bsub}$ compared to Ku_{Bsub} (Figure 2) suggests that the ability to bridge two DNA molecules together is not absolutely required to perform an efficient ligation, at least in our *in vitro* conditions. However, we cannot rule out that in germinating spores where NHEJ is active

(50), the ability of Ku_{Bsub} to bridge DNA molecules could stimulate the ligase activity of LigD.

Based on the DNA bridging assay, Ku_{Bsub} could bridge two linear DNA molecules by their ends (end-to-end bridging) and/or by their internal parts (adjacent bridging). By TEM, we did not observe end-to-end bridges (demonstrated, for instance, by the visualization of DNA molecules bridged by their ends with a Ku nucleofilament located at the junction). However, we have visualized molecules interacting by their end-proximal internal parts covered by Ku_{Bsub} (Supplementary Figure S7 panels A1, B and E). These observations favor an adjacent bridging mechanism dependent of the Ku_{Bsub} extended Cter domain which could be a primary step toward an end-to-end repositioning required for an efficient ligation. These DNA bridging sequential events have been proposed for the human NHEJ machinery *in vivo* (51). Finally, the human Ku heterodimer binds DNA with a defined polarity, Ku70 and Ku80 being located respectively proximal and distal to the DNA end (21,52). It is currently not known whether the bacterial Ku protein binds DNA ends in a preferred orientation. However, if the C-terminal domain played a crucial role in determining binding polarity, it is likely that the Ku_{Bsub} C-terminal domain mutants would have been affected in their threading abilities. The similar stability of the nucleoprotein filaments formed by Ku_{Bsub} and the C-terminal mutants as well as their conserved ability to thread is not in favor of such a role for the C-terminal domain of Ku_{Bsub}.

The ubiquitous presence of the minimal Cter region in the prokaryotic Ku proteins (Figure 1) strongly suggests that its role in the Ku_{Bsub}–$LigD_{Bsub}$ interaction is evolutionary conserved. The slight increase of the stimulation of $LigD_{Paer}$ using a Ku_{Paer} mutant truncated of its extended Cter region compared to a mutant similar to the Ku core protein used here (7) shows that the minimal Cter of Ku_{Paer} is also involved in the stimulation of $LigD_{Paer}$. The fact that this domain is important for the $LigD_{Bsub}$ and $LigD_{Paer}$ stimulation (Figure 2 and (7)) as well as for the formation of the Ku_{Bsub}–$LigD_{Bsub}$ complex (Figure 4) prompt us to propose that its role in promoting the Ku–LigD interaction is conserved in prokaryotes. The segment encompassing residues 111 to 273 of Ku_{Mtub} was sufficient to interact with $LigD_{Mtub}$ in a yeast two-hybrid assay (28). This segment contains the minimal Cter region (see Figure 1) which is in agreement with our hypothesis.

The prevalent addition of the basic extended tail to the C-terminus of bacterial Ku proteins suggests also that its role is conserved. This lysine-rich C-terminal sub-region is involved in the interaction of Ku_{Paer}, Ku_{Msme} and Ku_{Bsub} with DNA and its basic property suggests that it interacts with this molecule (7,27). We demonstrate this hypothesis by showing that addition of the extended Cter of Ku_{Bsub} to the GFP protein allows the resulting GFP-KC chimera to interact with DNA (Figure 5). Moreover, we show by electron microscopy that this extended domain limits the DNA threading ability of Ku_{Bsub} (Figure 8). Interestingly, Ku_{Mtub} does not contain a basic extended Cter region and it was proposed to translocate along linear dsDNA molecule (6). The ability of Ku_{Msme} for inward translocation from DNA ends remains to be investigated but a Ku_{Msme} truncated for its lysine rich C-terminal domain is able to translocate along a linear DNA molecule (27). Finally, the threading ability of Ku_{Paer} was not visualized by EMSA experiment using a 2.7 kbp linear dsDNA (7). However, using the same methodology, we were able to observe multiple discrete retarded bands (corresponding certainly to the threading of Ku_{Bsub} inward the DNA molecules as demonstrated by TEM in Figure 6) only with linear dsDNA substrate lesser than 500 bp length (not shown). Then, most of the bacterial Ku proteins, if not all, are able to thread inward a linear dsDNA molecule as their eukaryotic Ku70/80 homologues and we propose that the addition of a basic extended Cter domain limits (or prevents) this translocation mechanism. While the human Ku protein has been shown to bridge two DNA molecules *in vitro* (53), we discovered that a bacterial Ku protein also exhibited this activity. Future studies will be needed to determine whether the bridging activity is conserved across bacterial Ku proteins.

The threading ability of eukaryotic Ku proteins is proposed to increase the efficiency of the NHEJ mechanism by allowing the binding of other NHEJ factors at the DNA ends (25) and by displacing other proteins bound to the DNA such as histones (26). Similar roles can be hypothesized for the threading of Ku_{Bsub} on DNA, i.e. allowing the binding of $LigD_{Bsub}$ and its accessibility to DNA ends and displacement of other proteins bound in the vicinity of the DSB. Since bacterial Ku (at least Ku_{Bsub} and Ku_{Paer}) display a 5′-dRP lyase activity on abasic site (17) as the human Ku70/80 (16), translocation inside the DNA molecule could also increase the repair of abasic site surrounding a DSB. Finally, dsDNA protection against exonuclease and endonuclease activities *in vitro* by Ku_{Bsub} (Figure 9) suggests that Ku could protect the DNA ends of a broken chromosome against degradation by cellular enzymes. So, why did *B. subtilis* (and probably most of the NHEJ capable bacteria) conserve a domain limiting this threading ability?

Although no data supports a putative toxic effect of threading, one can imagine that translocation of Ku inside a broken chromosome could affect the genome integrity by displacing useful DNA repair proteins or by creating roadblocks against other cellular machineries. Thus mechanisms to regulate threading should exist. Eukaryotic Ku70s harbor a SAP domain at their C-terminal ends. This domain is thought to interact with the DNA (54) and could be important for the control of the threading ability of the Ku heterodimer (21). This view is consistent with structural changes of this extension observed upon DNA binding in the human Ku heterodimer, suggesting a direct interaction between the DNA molecule and this SAP domain (55,56). Also, the binding regulation of the yeast Ku to DNA could involve the sumoylation of the Ku70 C-terminal domain (57). However, the absence of a Ku mutant with an altered translocating ability does not allow demonstrating this threading regulation hypothesis. Besides this putative regulating SAP domain, it has been shown that the threading ability of the human Ku70/80 inside the DNA molecule is finely controlled by the DNA–PK catalytic subunit (58), strongly suggesting that ways to regulate the threading mechanism are required to allow an efficient NHEJ process and/or preventing toxic effect in eukaryotes. Prokaryotic Kus, with the exception of one Ku from *Streptomyces coelicolor*, do not contain the SAP domain (19,20)

and no prokaryotic homologue of DNA–PKcs has been identified. Then, we propose that the extended Cter region of Ku_{Bsub}, and possibly of other bacterial Ku, has evolved to keep a Ku-threading regulating function, as it is probably the case for the eukaryotic SAP domain. The molecular mechanism by which this regulation occurs could be electrostatic interactions between this lysine rich tail and the phosphodiester backbone of a DNA molecule, as suggested for Ku_{Paer} (7). Such interaction between this domain and the encircled DNA molecule should limit the threading of Ku inward the DNA, a phenomenon amplified by an increase of the length of the thread. We cannot exclude that differences in length of nucleoprotein filaments formed by Ku_{Bsub} and the C-terminal deletion mutants could be due to altered on- and/or off-rate constants of binding to DNA ends in the mutants. However, two lines of evidence do not support such a hypothesis: (i) the number of retarded bands observed by EMSA with the different Ku proteins were similar at concentration allowing a limited number of binding events (3 at 25 nM of proteins, Figure 6A) and (ii) a similar stability of the filaments formed by these Ku proteins that protect DNA from endonuclease cleavage was observed (Figure 9D).

Finally, this Ku-threading regulating role of the extended Cter as well as its role in the bridging of DNA molecules raise the question of the efficiency of the NHEJ pathway in bacteria encoding a Ku protein without this region, such as the Ku from *M. tuberculosis*, as well as the way to control the bridging and threading abilities of Ku in these bacteria.

SUPPLEMENTARY DATA

Supplementary Data are available at NAR Online.

ACKNOWLEDGEMENTS

We would like to thank P. Setlow, O. Delumeau, E. Dervyn and M. Ventroux for insightful discussions and Alex Dajkovic for critical reading of the manuscript. We are very grateful to the reviewers for their comments and useful suggestions.
Access to SOLEIL was granted by the Beam Allocation Group 'Plant and animal proteins: application in health, green chemistry and food application'. SAXS data were collected on beamline SWING at Synchrotron SOLEIL (Gif-sur-Yvette, France).

FUNDING

European Commission 7th Framework project BaSynthec [FP7-244093]; J.B.C. and E.L.C. acknowledge l'Agence Nationale de la Recherche [ANR Grants NHEJ-complexes N° ANR-12-BSV8-0012-03] for financial support. Funding for open access charge: European Commission 7th Framework project BaSynthec [FP7-244093].
Conflict of interest statement. None declared.

REFERENCES

1. Mehta,A. and Haber,J.E. (2014) Sources of DNA double-strand breaks and models of recombinational DNA repair. *Cold Spring Harb. Perspect. Biol.*, **6**, a016428.
2. Bennett,C.B., Lewis,A.L., Baldwin,K.K. and Resnick,M.A. (1993) Lethality induced by a single site-specific double-strand break in a dispensable yeast plasmid. *Proc. Natl. Acad. Sci. U.S.A.*, **90**, 5613–5617.
3. Moynahan,M.E. and Jasin,M. (2010) Mitotic homologous recombination maintains genomic stability and suppresses tumorigenesis. *Nat. Rev. Mol. Cell Biol.*, **11**, 196–207.
4. Aparicio,T., Baer,R. and Gautier,J. (2014) DNA double-strand break repair pathway choice and cancer. *DNA Repair (Amst)*, **19**, 169–175.
5. Deriano,L. and Roth,D.B. (2013) Modernizing the nonhomologous end-joining repertoire: alternative and classical NHEJ share the stage. *Annu. Rev. Genet.*, **47**, 433–455.
6. Weller,G.R., Kysela,B., Roy,R., Tonkin,L.M., Scanlan,E., Della,M., Devine,S.K., Day,J.P., Wilkinson,A., d'Adda di Fagagna,F. *et al.* (2002) Identification of a DNA nonhomologous end-joining complex in bacteria. *Science*, **297**, 1686–1689.
7. Zhu,H. and Shuman,S. (2010) Gap filling activities of *Pseudomonas* DNA ligase D (LigD) polymerase and functional interactions of LigD with the DNA end-binding Ku protein. *J. Biol. Chem.*, **285**, 4815–4825.
8. de Vega,M. (2013) The minimal *Bacillus subtilis* nonhomologous end joining repair machinery. *PLoS One*, **8**, e64232.
9. Bartlett,E.J., Brissett,N.C. and Doherty,A.J. (2013) Ribonucleolytic resection is required for repair of strand displaced nonhomologous end-joining intermediates. *Proc. Natl. Acad. Sci. U.S.A.*, **110**, E1984–E1991.
10. Della,M., Palmbos,P.L., Tseng,H.M., Tonkin,L.M., Daley,J.M., Topper,L.M., Pitcher,R.S., Tomkinson,A.E., Wilson,T.E. and Doherty,A.J. (2004) Mycobacterial Ku and ligase proteins constitute a two-component NHEJ repair machine. *Science*, **306**, 683–685.
11. Zhu,H. and Shuman,S. (2005) Novel 3′-ribonuclease and 3′-phosphatase activities of the bacterial non-homologous end-joining protein, DNA ligase D. *J. Biol. Chem.*, **280**, 25973–25981.
12. Aniukwu,J., Glickman,M.S. and Shuman,S. (2008) The pathways and outcomes of mycobacterial NHEJ depend on the structure of the broken DNA ends. *Genes Dev.*, **22**, 512–527.
13. Bhattarai,H., Gupta,R. and Glickman,M.S. (2014) DNA ligase C1 mediates the LigD-independent nonhomologous end-joining pathway of *Mycobacterium smegmatis*. *J. Bacteriol.*, **196**, 3366–3376.
14. Sinha,K.M., Stephanou,N.C., Gao,F., Glickman,M.S. and Shuman,S. (2007) Mycobacterial UvrD1 is a Ku-dependent DNA helicase that plays a role in multiple DNA repair events, including double-strand break repair. *J. Biol. Chem.*, **282**, 15114–15125.
15. Li,Z., Wen,J., Lin,Y., Wang,S., Xue,P., Zhang,Z., Zhou,Y., Wang,X., Sui,L., Bi,L.J. *et al.* (2011) A Sir2-like protein participates in mycobacterial NHEJ. *PLoS One*, **6**, e20045.
16. Roberts,S.A., Strande,N., Burkhalter,M.D., Strom,C., Havener,J.M., Hasty,P. and Ramsden,D.A. (2010) Ku is a 5′-dRP/AP lyase that excises nucleotide damage near broken ends. *Nature*, **464**, 1214–1217.
17. de Ory,A., Zafra,O. and de Vega,M. (2014) Efficient processing of abasic sites by bacterial nonhomologous end-joining Ku proteins. *Nucleic Acids Res.*, **42**, 13082–13095.
18. Grundy,G.J., Moulding,H.A., Caldecott,K.W. and Rulten,S.L. (2014) One ring to bring them all–the role of Ku in mammalian non-homologous end joining. *DNA Repair (Amst)*, **17**, 30–38.
19. Aravind,L. and Koonin,E.V. (2001) Prokaryotic homologs of the eukaryotic DNA-end-binding protein Ku, novel domains in the Ku protein and prediction of a prokaryotic double-strand break repair system. *Genome Res.*, **11**, 1365–1374.
20. Doherty,A.J., Jackson,S.P. and Weller,G.R. (2001) Identification of bacterial homologues of the Ku DNA repair proteins. *FEBS Lett.*, **500**, 186–188.
21. Walker,J.R., Corpina,R.A. and Goldberg,J. (2001) Structure of the Ku heterodimer bound to DNA and its implications for double-strand break repair. *Nature*, **412**, 607–614.
22. de Vries,E., van Driel,W., Bergsma,W.G., Arnberg,A.C. and van der Vliet,P.C. (1989) HeLa nuclear protein recognizing DNA termini and translocating on DNA forming a regular DNA-multimeric protein complex. *J. Mol. Biol.*, **208**, 65–78.
23. Paillard,S. and Strauss,F. (1991) Analysis of the mechanism of interaction of simian Ku protein with DNA. *Nucleic Acids Res.*, **19**, 5619–5624.
24. Downs,J.A. and Jackson,S.P. (2004) A means to a DNA end: the many roles of Ku. *Nat. Rev. Mol. Cell Biol.*, **5**, 367–378.

25. Yoo,S. and Dynan,W.S. (1999) Geometry of a complex formed by double strand break repair proteins at a single DNA end: recruitment of DNA-PKcs induces inward translocation of Ku protein. *Nucleic Acids Res.*, **27**, 4679–4686.
26. Roberts,S.A. and Ramsden,D.A. (2007) Loading of the nonhomologous end joining factor, Ku, on protein-occluded DNA ends. *J. Biol. Chem.*, **282**, 10605–10613.
27. Kushwaha,A.K. and Grove,A. (2013) *Mycobacterium smegmatis* Ku binds DNA without free ends. *Biochem. J.*, **456**, 275–282.
28. Gong,C., Bongiorno,P., Martins,A., Stephanou,N.C., Zhu,H., Shuman,S. and Glickman,M.S. (2005) Mechanism of nonhomologous end-joining in mycobacteria: a low-fidelity repair system driven by Ku, ligase D and ligase C. *Nat. Struct. Mol. Biol.*, **12**, 304–312.
29. Pitcher,R.S., Brissett,N.C. and Doherty,A.J. (2007) Nonhomologous end-joining in bacteria: a microbial perspective. *Annu. Rev. Microbiol.*, **61**, 259–282.
30. Kushwaha,A.K. and Grove,A. (2013) C-terminal low-complexity sequence repeats of *Mycobacterium smegmatis* Ku modulate DNA binding. *Biosci. Rep.*, **33**, 175–184.
31. Casadaban,M.J. and Cohen,S.N. (1980) Analysis of gene control signals by DNA fusion and cloning in *Escherichia coli*. *J. Mol. Biol.*, **138**, 179–207.
32. Lewis,P.J. and Marston,A.L. (1999) GFP vectors for controlled expression and dual labelling of protein fusions in *Bacillus subtilis*. *Gene*, **227**, 101–110.
33. David,G. and Perez,J. (2009) Combined sampler robot and high-performance liquid chromatography: a fully automated system for biological small-angle X-ray scattering experiments at the Synchrotron SOLEIL SWING beamline. *J. Appl. Crystallogr.*, **42**, 892–900.
34. Konarev,P.V., Petoukhov,M.V., Volkov,V.V. and Svergun,D.I. (2006) ATSAS 2.1, a program package for small-angle scattering data analysis. *J. Appl. Crystallogr.*, **39**, 277–286.
35. Konarev,P.V., Volkov,V.V., Sokolova,A.V., Koch,M.H.J. and Svergun,D.I. (2003) PRIMUS: a Windows PC-based system for small-angle scattering data analysis. *J. Appl. Crystallogr.*, **36**, 1277–1282.
36. Svergun,D.I. (1992) Determination of the regularization parameter in indirect-transform methods using perceptual criteria. *J. Appl. Crystallogr.*, **25**, 495–503.
37. Fischer,H., de Oliveira Neto,M., Napolitano,H.B., Polikarpov,I. and Craievich,A.F. (2010) Determination of the molecular weight of proteins in solution from a single small-angle X-ray scattering measurement on a relative scale. *J. Appl. Crystallogr.*, **43**, 101–109.
38. Rambo,R.P. and Tainer,J.A. (2013) Accurate assessment of mass, models and resolution by small-angle scattering. *Nature*, **496**, 477–481.
39. Svergun,D.I., Petoukhov,M.V. and Koch,M.H. (2001) Determination of domain structure of proteins from X-ray solution scattering. *Biophys. J.*, **80**, 2946–2953.
40. Volkov,V.V. and Svergun,D.I. (2003) Uniqueness of ab initio shape determination in small-angle scattering. *J. Appl. Crystallogr.*, **36**, 860–864.
41. Kozin,M.B. and Svergun,D.I. (2001) Automated matching of high- and low-resolution structural models. *J. Appl. Crystallogr.*, **34**, 33–41.
42. Evrard,G., Mareuil,F., Bontems,F., Sizun,C. and Perez,J. (2011) DADIMODO: a program for refining the structure of multidomain proteins and complexes against small-angle scattering data and NMR-derived restraints. *J. Appl. Crystallogr.*, **44**, 1264–1271.
43. Svergun,D., Barberato,C. and Koch,M. H. J. (1995) CRYSOL - A program to evaluate x-ray solution scattering of biological macromolecules from atomic coordinates. *J. Appl. Crystallogr.*, **28**, 768–773.
44. Lutz,R. and Bujard,H. (1997) Independent and tight regulation of transcriptional units in *Escherichia coli* via the LacR/O, the TetR/O and AraC/I1-I2 regulatory elements. *Nucleic Acids Res.*, **25**, 1203–1210.
45. Dupaigne,P., Lavelle,C., Justome,A., Lafosse,S., Mirambeau,G., Lipinski,M., Pietrement,O. and Le Cam,E. (2008) Rad51 polymerization reveals a new chromatin remodeling mechanism. *PLoS One*, **3**, e3643.
46. Roy,S., Andres,S.N., Vergnes,A., Neal,J.A., Xu,Y., Yu,Y., Lees-Miller,S.P., Junop,M., Modesti,M. and Meek,K. (2012) XRCC4's interaction with XLF is required for coding (but not signal) end joining. *Nucleic Acids Res.*, **40**, 1684–1694.
47. Kobayashi,H., Simmons,L.A., Yuan,D.S., Broughton,W.J. and Walker,G.C. (2008) Multiple Ku orthologues mediate DNA non-homologous end-joining in the free-living form and during chronic infection of *Sinorhizobium meliloti*. *Mol. Microbiol.*, **67**, 350–363.
48. Nicolas,P., Mader,U., Dervyn,E., Rochat,T., Leduc,A., Pigeonneau,N., Bidnenko,E., Marchadier,E., Hoebeke,M., Aymerich,S. *et al.* (2012) Condition-dependent transcriptome reveals high-level regulatory architecture in *Bacillus subtilis*. *Science*, **335**, 1103–1106.
49. Ono,M., Tucker,P.W. and Capra,J.D. (1994) Production and characterization of recombinant human Ku antigen. *Nucleic Acids Res.*, **22**, 3918–3924.
50. Wang,S.T., Setlow,B., Conlon,E.M., Lyon,J.L., Imamura,D., Sato,T., Setlow,P., Losick,R. and Eichenberger,P. (2006) The forespore line of gene expression in *Bacillus subtilis*. *J. Mol. Biol.*, **358**, 16–37.
51. Reid,D.A., Keegan,S., Leo-Macias,A., Watanabe,G., Strande,N.T., Chang,H.H., Oksuz,B.A., Fenyo,D., Lieber,M.R., Ramsden,D.A. *et al.* (2015) Organization and dynamics of the nonhomologous end-joining machinery during DNA double-strand break repair. *Proc. Natl. Acad. Sci. U.S.A.*, **112**, E2575–E2584.
52. Yoo,S., Kimzey,A. and Dynan,W.S. (1999) Photocross-linking of an oriented DNA repair complex. Ku bound at a single DNA end. *J. Biol. Chem.*, **274**, 20034–20039.
53. Ramsden,D.A. and Gellert,M. (1998) Ku protein stimulates DNA end joining by mammalian DNA ligases: a direct role for Ku in repair of DNA double-strand breaks. *EMBO J.*, **17**, 609–614.
54. Aravind,L. and Koonin,E.V. (2000) SAP - a putative DNA-binding motif involved in chromosomal organization. *Trends Biochem. Sci.*, **25**, 112–114.
55. Rivera-Calzada,A., Spagnolo,L., Pearl,L.H. and Llorca,O. (2007) Structural model of full-length human Ku70-Ku80 heterodimer and its recognition of DNA and DNA-PKcs. *EMBO Rep.*, **8**, 56–62.
56. Lehman,J.A., Hoelz,D.J. and Turchi,J.J. (2008) DNA-dependent conformational changes in the Ku heterodimer. *Biochemistry*, **47**, 4359–4368.
57. Hang,L.E., Lopez,C.R., Liu,X., Williams,J.M., Chung,I., Wei,L., Bertuch,A.A. and Zhao,X. (2014) Regulation of Ku-DNA association by Yku70 C-terminal tail and SUMO modification. *J. Biol. Chem.*, **289**, 10308–10317.
58. Calsou,P., Frit,P., Humbert,O., Muller,C., Chen,D.J. and Salles,B. (1999) The DNA-dependent protein kinase catalytic activity regulates DNA end processing by means of Ku entry into DNA. *J. Biol. Chem.*, **274**, 7848–7856.
59. Sievers,F., Wilm,A., Dineen,D., Gibson,T.J., Karplus,K., Li,W., Lopez,R., McWilliam,H., Remmert,M., Soding,J. *et al.* (2011) Fast, scalable generation of high-quality protein multiple sequence alignments using Clustal Omega. *Mol. Syst. Biol.*, **7**, 539.
60. Garnier,J., Gibrat,J.F. and Robson,B. (1996) GOR method for predicting protein secondary structure from amino acid sequence. *Methods Enzymol.*, **266**, 540–553.