



**HAL**  
open science

## **Bacterial NHEJ: A never ending story**

Claire Bertrand, Annabelle Thibessard, Claude Bruand, François Lecointe,  
Pierre Leblond

► **To cite this version:**

Claire Bertrand, Annabelle Thibessard, Claude Bruand, François Lecointe, Pierre Leblond. Bacterial NHEJ: A never ending story. *Molecular Microbiology*, 2019, 10.1111/mmi.14218 . hal-02044826

**HAL Id: hal-02044826**

**<https://hal.univ-lorraine.fr/hal-02044826>**

Submitted on 21 Feb 2019

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

1 **Bacterial NHEJ: A never ending story**

2 Claire Bertrand<sup>1</sup>, Annabelle Thibessard<sup>1</sup>, Claude Bruand<sup>2\*</sup>, François Lecointe<sup>3\*</sup>, Pierre Leblond<sup>1\*</sup>

3 <sup>1</sup>Université de Lorraine, Inra, DynAMic, F-54000 Nancy, France

4

5 <sup>2</sup>Laboratoire des Interactions Plantes-Microorganismes, Université de Toulouse, INRA,  
6 CNRS, Castanet-Tolosan, France.

7

8 <sup>3</sup>Micalis Institute, INRA, AgroParisTech, Université Paris-Saclay, 78350 Jouy-en-Josas,  
9 France.

10

11 \*Co-corresponding authors

12

13

14 Running title: **Growing complexity of bacterial NHEJ**

15 Key words: **Non-Homologous End Joining, *Streptomyces*, *Sinorhizobium meliloti*, *Bacillus subtilis*,**

16 ***Mycobacterium***

17

18

19

20 **Abstract**

21 Double-strand breaks (DSBs) are the most detrimental DNA damage encountered by bacterial  
22 cells. DSBs can be repaired by homologous recombination thanks to the availability of an intact  
23 DNA template or by Non-Homologous End Joining (NHEJ) when no intact template is available.  
24 Bacterial NHEJ is performed by sets of proteins of growing complexity from *Bacillus subtilis* and  
25 *Mycobacterium tuberculosis* to *Streptomyces* and *Sinorhizobium meliloti*. Here we discuss the  
26 contribution of these models to the understanding of the bacterial NHEJ repair mechanism as  
27 well as the involvement of NHEJ partners in other DNA repair pathways. The importance of NHEJ  
28 and of its complexity is discussed in the perspective of regulation through the biological cycle of  
29 the bacteria and in response to environmental stimuli. Finally, we consider the role of NHEJ in  
30 genome evolution, notably in horizontal gene transfer.

## 31 I- Bacterial NHEJ: A two-component DNA repair mechanism?

32 Until *in silico* identification of putative NHEJ actors in bacterial genomes, it was thought that  
33 Homologous Recombination (HR) was the major double-strand break (DSB) repair mechanism in  
34 prokaryotes. However, HR needs an intact template to ensure repair of the break, a condition  
35 encountered by bacteria only during their dividing state and raising the question of the repair of  
36 DSB in other phases of their life cycle. The identification of a bacterial NHEJ answered, at least in  
37 part, this question since, in eukaryotes, NHEJ was known to ligate DNA ends together without  
38 any homologous template. This mechanism has been well documented in eukaryotes and  
39 involves a dozen of actors in human, including the Ku70/Ku80 heterodimer, and the ligase IV  
40 (Chang *et al.*, 2017).

41 The first hint for the existence of a bacterial NHEJ was provided by *in silico* approaches leading  
42 to the identification of Ku70/80 homologs and ATP-dependent DNA-ligase homologs (Doherty *et*  
43 *al.*, 2001; Aravind and Koonin, 2001). These Ku and DNA-ligase encoding genes are often  
44 colocalized, or in an operonic structure. The DNA-ligase genes encode a putative ligase domain  
45 (LigDom or LIG), fused or not with a polymerase domain (PolDom, POL or Prim-Pol) and/or a  
46 nuclease domain (NucDom or PE for Phosphoesterase). Whatever the domain combination,  
47 these multifunctional ligases were collectively named LigD, in reference to the main NHEJ ligase  
48 in Mycobacteria (Gong *et al.*, 2005).

49 The first experimental validation of a bacterial NHEJ came from the study of Weller *et al.* showing  
50 that *in vitro* the *Mycobacterium tuberculosis* LigD protein (LigD<sub>Mtub</sub>) is indeed an ATP-dependent  
51 DNA ligase that is stimulated by its cognate Ku<sub>Mtub</sub> partner, possibly via an interaction between  
52 Ku<sub>Mtub</sub> bound to DNA and LigD<sub>Mtub</sub> (Weller *et al.*, 2002). In parallel, authors described the

53 deleterious effect of the deletion of the *Bacillus subtilis* genes coding for LigD<sub>Bsub</sub> and Ku<sub>Bsub</sub>  
54 (*ykoU-ykoV* operon) on the survival of cells in stationary phase after ionizing radiation (IR). IR is  
55 a cause of multiple DNA lesions, including DSBs that are particularly hazardous. These two  
56 proteins from *B. subtilis* have been later shown to also perform NHEJ repair *in vitro* (de Vega,  
57 2013). The Ku and/or LigD-dependent ability of *M. tuberculosis* and *Mycobacterium smegmatis*  
58 to recircularize linear plasmid DNA after transformation, and the IR sensitivity of *M. smegmatis*  
59 *ku* and *ligD* mutants in stationary phase, have given further evidence *in vivo* that several bacterial  
60 species are NHEJ proficient (Stephanou *et al.*, 2007).

61 LigD<sub>Mtub</sub> is not a common ligase. Its PolDom provides LigD<sub>Mtub</sub> with DNA-dependent DNA  
62 polymerase, DNA-dependent RNA polymerase and primase and non-templated nucleotide  
63 addition activities (Della *et al.*, 2004). Moreover, the Poldom interacts *in vitro* with Ku<sub>Mtub</sub>-DNA  
64 complex suggesting that it could be responsible for the recruitment of LigD<sub>Mtub</sub> on Ku-bound DSB  
65 *in vivo* (Pitcher *et al.*, 2005). LigD<sub>Mtub</sub> also shows a 3'-5' exonuclease activity most probably carried  
66 by its NucDom (Della *et al.*, 2004). LigD from *Pseudomonas aeruginosa* (LigD<sub>Paer</sub>) and  
67 *Agrobacterium tumefaciens* (LigD2) display similar activities accordingly to their identical domain  
68 composition, though different domain organization, compared to LigD<sub>Mtub</sub> (Zhu and Shuman,  
69 2005a; Zhu and Shuman, 2005b; Zhu *et al.*, 2005; Zhu and Shuman, 2007). Thus, LigD displays all  
70 activities allowing to process DNA ends, if necessary, and ligate them without the need for other  
71 proteins, except Ku. Accordingly, expression of Ku and LigD from *M. tuberculosis* *i*) partially  
72 restores NHEJ in yeast *ku70* or *dnl4* mutant strains (Della *et al.*, 2004) and *ii*) renders *Escherichia*  
73 *coli* NHEJ-proficient (Malyarchuk *et al.*, 2007).

74 Altogether, these studies have led to the following model for the bacterial NHEJ: Ku homodimer  
75 binds to DNA ends generated by a break and recruits LigD (one multidomain polypeptide, such

76 as LigD<sub>Mtub</sub>, or a holoenzyme consisting of several subunits similar to the LigDom, PolDom and/or  
77 NucDom of LigD<sub>Mtub</sub>) probably via a direct physical interaction on the DNA between Ku and the  
78 PolDom of the LigD holoenzyme. LigD is able to process DNA ends, if broken ends are non-directly  
79 ligatable, thanks to its PolDom and/or NucDom, before ligation by LigDom (Pitcher *et al.*, 2007a).

80 Most of the past studies on bacterial NHEJ have already been extensively reviewed (Bowater and  
81 Doherty, 2006; Pitcher *et al.*, 2007a; Shuman and Glickman, 2007; Gu and Lieber, 2008; Brissett  
82 and Doherty, 2009), with the exception of the most recent ones on structural data of different  
83 LigD domains and mechanistic details of DNA end processing by NHEJ actors before ligation (Nair  
84 *et al.*, 2010; Zhu and Shuman, 2010; Brissett *et al.*, 2011;; Brissett *et al.*, 2013). The archaea  
85 *Methanocella paludicola* contain genes coding for homologs of PolDom, NucDom, LigDom and  
86 Ku. Their characterization have also deepened our knowledge on the processing of DNA ends  
87 during NHEJ (Bartlett *et al.*, 2013; Brissett *et al.*, 2013; Bartlett *et al.*, 2016). In the present report,  
88 we wished to focus on recent data obtained with different bacterial models highlighting the  
89 complexity of the prokaryotic NHEJ pathway in terms of involved actors, their interactions, their  
90 regulation, their links with other DNA repair pathways and finally their impact on bacterial  
91 evolution.

## 92 **II- Multiple components NHEJ pathways**

93 The above model of minimal NHEJ, with one Ku and one LigD holoenzyme performing bacterial  
94 NHEJ, does not take into account the existence of other proteins involved in this mechanism.  
95 These proteins can be orthologues of Ku or of domains of the LigD holoenzyme, and/or proteins  
96 interacting with them. Figure 1 illustrates diverse levels of NHEJ complexity in four bacterial  
97 models.

98 **NHEJ complexity: multiplicity of NHEJ-like genes**

99 A recent and large scale search for *ku*-like genes suggests the existence of a potential NHEJ repair  
100 system in about 25% of the prokaryotes for which a genome was available in public databases  
101 (McGovern *et al.*, 2016), with a scattered distribution. Among the bacterial species predicted to  
102 possess a NHEJ repair system, some actually encode several putative Ku homologs, mainly  
103 actinobacteria of the streptomycetes family, and proteobacteria of the  $\alpha$  subdivision (McGovern  
104 *et al.*, 2016). For instance, more than 70% of *Streptomyces* species encode more than one *ku*  
105 gene (up to 4 in *Streptomyces coelicolor*) (Hoff *et al.*, 2016), and more than half of  $\alpha$ -  
106 proteobacteria possessing a *ku* gene actually have two or more (up to 6 in *Rhizobium*  
107 *leguminosarum bv trifolii*). Some of the *ku* genes are located on plasmids, suggesting their  
108 putative acquisition via horizontal gene transfer. Importance of multiple *ku* genes in some NHEJ  
109 proficient bacteria is discussed further below.

110 Putative NHEJ ligases are also often present in multiple copies. However, because of the more  
111 variable organization of LigD, it is more difficult in this case to predict whether the genes encode  
112 *bona fide* NHEJ ligases. This difficulty is reinforced by the fact that several ATP-dependent DNA  
113 ligases can be used to perform NHEJ in the same bacterium. For instance, *M. smegmatis* encodes  
114 four ATP-dependent DNA ligases (LigB<sub>Msm</sub>, LigC1<sub>Msm</sub>, LigC2<sub>Msm</sub> and LigD<sub>Msm</sub>) (Gong *et al.*, 2005).  
115 Among them, LigD<sub>Msm</sub> is the main NHEJ ligase and contains the PolDom, LigDom and NucDom  
116 domains, as its homolog LigD<sub>Mtub</sub>. However, LigC1<sub>Msm</sub> (a LigDom-only enzyme) can act with Ku in  
117 an alternative NHEJ in *M. smegmatis* (Gong *et al.*, 2005; Akey *et al.*, 2006; Aniukwu *et al.*, 2008;  
118 Bhattarai *et al.*, 2014). This LigC1<sub>Msm</sub>-dependent alternative NHEJ requires the PolDom of  
119 LigD<sub>Msm</sub>, the LigDom and NucDom being dispensable for this mechanism (Bhattarai *et al.*, 2014).  
120 This result, supplemented by the fact that the PolDom of LigD<sub>Msm</sub> plays a structural role

121 independent of its polymerase activity in the efficiency of recircularization of linearized plasmid  
122 *in vivo* (Aniukwu *et al.*, 2008), highlights that the PolDom could serve as an interaction platform  
123 for an ATP-dependent ligase protein containing only a LigDom. Other LigC proteins, LigC1 and  
124 LigC2 from *A. tumefaciens* are endowed with ligase activities that are stimulated by Ku (Zhu and  
125 Shuman, 2007), suggesting that alternative NHEJ pathways are conserved across bacterial  
126 species. It has to be noticed that *M. smegmatis* contains two *polD* genes expressing two PolDom-  
127 only polymerases that can suppress partially the sensitivity to IR of a mutant encoding a  
128 polymerase deficient LigD, suggesting that, as observed with LigC1, several PolDom proteins can  
129 be used in the NHEJ pathway in Mycobacteria (Bhattarai *et al.*, 2014).

130 In *Streptomyces*, PolDom, LigDom and NucDom only exist as stand-alone domains and not as  
131 multifunctional proteins, making even more complex the identification of actual NHEJ partners.  
132 The analysis of 38 *Streptomyces* genomes revealed a conserved gene set, composed of two loci,  
133 *kuA* and its neighbor *polK* and *ligC* (encoding a putative ligase displaying an higher level of  
134 homology with LigC1 from *M. smegmatis* compared to LigD<sub>M<sub>sme</sub></sub>) and its neighbor *polR* (*polK* and  
135 *polR* being two PolDom encoding genes). Surprisingly, *ligD*, whose product is the closest homolog  
136 of LigD<sub>M<sub>tub</sub></sub>, is not ubiquitous in these genomes (present in 75% of the analyzed species),  
137 suggesting that the core *ligC* gene provides the main NHEJ ligase function. The variable set of  
138 *Streptomyces* NHEJ-like genes also includes *kuB* and *kuC*, the latter often being plasmid borne  
139 and often laying in the close vicinity of *ligD*. Although they are not conserved, these genes take  
140 part in DNA repair since their mutation in *Streptomyces ambofaciens* confers spore sensitivity to  
141 IR and increases the sensitivity of the core gene set mutants (Hoff *et al.*, 2016). In *S. ambofaciens*,  
142 the involvement of KuA in a NHEJ system has been demonstrated together with the existence of



143 a secondary illegitimate DSB repair system that could imply these variable genes (Hoff *et al.*,  
144 2018).

145 The *Streptomyces* organization also questions the role of the two conserved PolDom encoding  
146 genes (*polK* and *polR*). In this regard, *Streptomyces* is not an exception since *M. smegmatis* (see  
147 above and below, PolD1, PolD2) and numerous eukaryotes, especially vertebrates, also possess  
148 several polymerases involved in NHEJ (Ramsden and Asagoshi, 2012). A first hypothesis to explain  
149 this observation is based on the difference in expression profile: indeed, *S. ambofaciens polR*  
150 seems constitutively expressed whereas *polK* transcription was not detected in the tested  
151 conditions (Hoff *et al.*, 2016), but might be induced in response to some genotoxic stress.  
152 Alternatively or concomitantly, these PolDoms could display different template requirements, as  
153 for the four polymerases of the X family found in vertebrates (Pol $\beta$ , Pol $\lambda$ , Pol $\mu$  and TdT, for review  
154 see Ramsden and Asagoshi, 2012). Therefore, each PolDom could act specifically on different  
155 types of damaged DNA ends. In favor of this hypothesis, in *M. smegmatis*, a mutation of the  
156 PolDom of LigD abolishes the nucleotide addition but not the fill-in activity suggesting that at  
157 least two polymerase activities (probably driven by PolD1 and/or PolD2, Bhattarai *et al.*, 2014)  
158 would be required for full DSB repair achievement (Aniukwu *et al.*, 2008). Finally, due to its  
159 similarity with the mycobacterial locus encoding *ligC1* and *prim-polC* involved in a BER system  
160 (Płociński *et al.*, 2017), we surmise that the *Streptomyces ligC-polR* locus could also be involved  
161 in this DNA repair pathway in addition to its role in NHEJ.

162 Another bacterium displaying a complex organization is *Sinorhizobium meliloti*, an  $\alpha$ -  
163 proteobacterium which encodes four Ku and four LigD orthologues (Galibert *et al.*, 2001); here,  
164 LigD proteins are defined as multidomain ATP-dependent ligases, with LigD3 and LigD1 lacking a  
165 Nuc and a Pol domain, respectively. The Ku proteins of  $\alpha$ -proteobacteria can be classified in

166 several families based on sequence similarity, but a “representative” Ku protein present in all  $\alpha$ -  
167 proteobacteria could not be identified. Single and multiple *S. meliloti* *ku* mutations conferred IR  
168 sensitivity to stationary phase bacteria, suggesting a role of these genes in DNA repair (Kobayashi  
169 *et al.*, 2008; Dupuy *et al.*, 2017). Involvement of the various *ku* and *ligD* genes in NHEJ was  
170 assessed more precisely using a linear plasmid recircularization assay, which showed the  
171 existence of at least two distinct NHEJ pathways (Fig. 1) (Dupuy *et al.*, 2018). The first major  
172 pathway is dependent on *ku2* and *ligD2*, two adjacent genes apparently organized in operon,  
173 although they can be transcribed from independent promoters. A second, less efficient pathway,  
174 dependent on *ku3*, *ku4* and *ligD4* was also detected in stress conditions (see further below). Both  
175 *ku3* and *ku4* are organized in operon, whereas *ligD4* is located elsewhere in the genome, but all  
176 three genes are co-regulated (see further below) and *ku3*, *ku4* and *ligD4* orthologues are often  
177 found in close proximity, mostly in operon, in  $\alpha$ -proteobacterial genomes. Interestingly, genetic  
178 analyses showed that both *ku3* and *ku4* are needed for activity of this second pathway. This  
179 observation, together with the quasi-systematic co-occurrence and co-regulation of *ku3*- and  
180 *ku4*-like genes in NHEJ equipped  $\alpha$ -proteobacteria, led to the suggestion that Ku3 and Ku4 might  
181 function as heterodimers, like eukaryotic Ku proteins. Although LigD4 was systematically found  
182 associated to the Ku3-Ku4-dependent pathway, LigD2, and to a much lesser extent LigD3, are  
183 also able to function, probably aspecifically, with Ku3 and Ku4 on certain types of DNA ends and  
184 in certain genetic contexts. Both *ku1* and *ligD1* are weakly expressed, and consistently, no NHEJ  
185 activity could be associated with these genes in the tested conditions. PolD1 (also called LigD1p,  
186 Kobayashi *et al.*, 2008), encoded by the gene located just downstream *ligD1*, is likely the vestigial  
187 PolDom of a larger ancestral LigD1 protein. Deletions of *ligD1* analyzed by Dupuy *et al.* (2018)  
188 encompassed *polD1*, showing that this gene is not involved in NHEJ in the tested conditions.

189 *S. meliloti* also encodes three *ligC* and one *ligB* genes, whose role in NHEJ or another DNA repair  
190 system has not been tested so far (Fig. 1).

191 Inactivation of NHEJ in *M. smegmatis* and *S. meliloti* unmasked a background level of Ku- and  
192 LigD-independent end-joining activity, called alternative End-Joining (A-EJ), observed during  
193 recircularization of linear transforming plasmids (Aniukwu *et al.*, 2008; Dupuy *et al.*, 2018).  
194 Depending on the bacterial species and the DNA end type, the observed efficiencies and fidelity  
195 of repair greatly varied. It is unclear whether A-EJ occurs by the same mechanism in all cases, but  
196 as it was observed in *M. smegmatis ligBCD* mutants, it is believed to depend on the replicative  
197 ligase LigA. Interestingly, LigA-dependent, LigB-independent A-EJ has been described in the NHEJ-  
198 deficient *E. coli* (Chayot *et al.*, 2010). When not resulting in faithful ligation of compatible DNA  
199 ends (which occurred in some cases with 100% efficiency; Aniukwu *et al.*, 2008; Dupuy *et al.*,  
200 2018), A-EJ repair events mostly involved deletions of variable lengths, with frequent occurrence  
201 of microhomologies at the junctions (Aniukwu *et al.*, 2008), very similar to the observed products  
202 of A-EJ in *E. coli* (Chayot *et al.*, 2010). These observations highlight the importance of DNA end  
203 resection for homology search, which in *E. coli* is in part due to RecBCD. The resection enzymes  
204 involved in the A-EJ of NHEJ-proficient bacteria have not been characterized yet.

#### 205 **NHEJ complexity: are there other actors involved?**

206 Proteins other than Ku and LigD domains have been suggested to be involved in NHEJ, either via  
207 their physical and/or functional interaction with Ku and/or LigD or via the impact of their  
208 inactivation on DSB repair. This is the case for the helicase UvrD1 from Mycobacteria, which  
209 interacts with Ku in a yeast two-hybrid assay and which forms ternary complex with Ku and  
210 dsDNA (Sinha *et al.*, 2007). The fact that the helicase activity of UvrD1 is stimulated in presence  
211 of Ku<sub>Msm</sub> and that the deletion of *uvrD1* in *M. smegmatis* sensitizes this bacterium to IR makes

212 sense with a role of UvrD1 in DSB repair via NHEJ. However, its role in this process remains to be  
213 elucidated.

214 Another protein interacting with the NHEJ actors in Mycobacteria is Sir2. This deacetylase  
215 interacts physically with both Ku and LigD and its inactivation decreased the NHEJ efficiency *in*  
216 *vivo* (Li *et al.*, 2011). Its putative role in the NHEJ process is discussed below (see “Regulation of  
217 NHEJ”).

### 218 **NHEJ complexity: the multiple DNA binding abilities of Ku**

219 Bacterial Ku is composed of a N-terminal Ku core domain homologous to the central domain of  
220 the Ku70 and Ku80 eukaryotic homologs (Aravind and Koonin, 2001; Kushwaha and Grove,  
221 2013a; McGovern *et al.*, 2016; Pitcher *et al.*, 2007a). Ku70/80 heterodimer has been shown to  
222 bind DSB and to form a ring (involving the Ku core domains of the two subunits) able to encircle  
223 a dsDNA molecule (Walker *et al.*, 2001). This particular structure of the eukaryotic Ku allows its  
224 threading inward the DNA molecule when additional Ku proteins enter by the ends (de Vries *et*  
225 *al.*, 1989; Paillard and Strauss, 1991; Downs and Jackson, 2004). This threading ability is  
226 conserved at least for Ku from *B. subtilis*, *M. tuberculosis* and most probably *M. smegmatis*  
227 (Weller *et al.*, 2002; Kushwaha and Grove, 2013a; McGovern *et al.*, 2016). The Ku core domain of  
228 bacterial Ku is associated to a variable C-terminal (C-ter) tail which contains the ubiquitous  
229 minimal C-ter domain and, in most cases, an extended C-ter tail (McGovern *et al.*, 2016). The  
230 minimal C-ter is required for the interaction of Ku<sub>Bsub</sub> with LigD<sub>Bsub</sub>, is involved in the stimulation  
231 of the ligase activity of LigD<sub>Bsub</sub> and LigD<sub>Paer</sub> and is contained in a segment of Ku<sub>Mtub</sub> interacting  
232 with LigD<sub>Mtub</sub> in a yeast two-hybrid assay (Gong *et al.*, 2005; Zhu and Shuman, 2010; McGovern  
233 *et al.*, 2016). These observations led to the proposal that the minimal C-ter is a part of the LigD-

234 interacting domain in bacterial Ku. The extended C-ter is predominantly a domain rich in basic  
235 residues (lysine and arginine) that is involved, but not required, in the Ku-dependent stimulation  
236 of the ligase activity of LigD in *P. aeruginosa* and *B. subtilis* (Zhu and Shuman, 2010; McGovern *et*  
237 *al.*, 2016). This domain also confers to Ku the ability to bind DNA without free ends (Kushwaha  
238 and Grove, 2013b; McGovern *et al.*, 2016) and limits the threading of Ku<sub>Bsub</sub> inward the DNA  
239 molecule. Its function could be to reduce putative negative effects of this threading inside the  
240 chromosome on other DNA repair processes. The regulation of the threading of Ku may also be  
241 important for the recruitment of LigD near, or at the DSB site, ensuring an efficient NHEJ repair  
242 since LigD alone is a poorly efficient ligase without its Ku partner *in vitro* (Weller *et al.*, 2002; Della  
243 *et al.*, 2004; Zhu and Shuman, 2010; de Ory *et al.*, 2014; McGovern *et al.*, 2016). The function of  
244 Ku in the NHEJ could be more proactive than previously anticipated by allowing the bridging of  
245 DNA ends, thanks to its extended C-ter (McGovern *et al.*, 2016), an ability shared by the human  
246 Ku (Ramsden and Gellert, 1998). The different functions of this extended C-ter domain raise the  
247 question about the efficiency of the NHEJ in bacteria encoding a Ku without this domain and/or  
248 about the role of Ku and its extended C-ter in another DNA repair mechanism (for instance the  
249 BER process, see below). In *Streptomyces*, one of the accessory Ku (KuB), present in only 8 of 38  
250 species (Hoff *et al.*, 2016), possesses an extra C-ter domain weakly related to the RhoN domain  
251 (Pfam 07498) which role in Ku protein remains to be elucidated.

252 Finally, a very recent study demonstrated that the eukaryotic Ku orthologue encoded by the  
253 bacteriophage Mu (d'Adda di Fagagna *et al.*, 2003) is able to promote NHEJ repair of plasmidic  
254 and chromosomal DSBs, most probably by interacting with and stimulating the *E. coli* replicative  
255 ligase LigA (Bhattacharyya *et al.*, 2018). Surprisingly, MuGam also stimulates *in vitro* the LigB  
256 ligase of *E. coli*, although MuGam and LigB do not seem to interact on DNA. These results raise

257 the possibility that a bacterial Ku protein might also increase the ligase activity of a non-NHEJ  
258 ligase. In this line, Ku<sub>Msmeg</sub> (Kushwaha and Grove, 2013a), but not Ku<sub>Mtub</sub> (Weller *et al.*, 2002), has  
259 been shown to stimulate the T4 DNA ligase *in vitro*. To our knowledge, no study assessed the  
260 possibility that a bacterial Ku redirects the activity of a bacterial replicative ligase.

### 261 III- Regulation of NHEJ

262 In eukaryotes, NHEJ repairs DSBs throughout the whole cell cycle and is predominant in G1 phase  
263 while HR relies on the presence of an intact sister chromatid and therefore functions in S/G2  
264 phase (Ira *et al.*, 2004; Aylon *et al.*, 2004; Clerici *et al.*, 2008). Similarly, during quiescent states  
265 such as sporulation (*Bacillus*, *Streptomyces*) or late-stationary phase (*Mycobacterium*,  
266 *Sinorhizobium*) in bacteria, HR would not be operational because stationary phase cells and  
267 spores are assumed to be single genome compartments. Note however that spores (Setlow and  
268 Setlow, 1996; Moeller *et al.*, 2007; Hoff *et al.*, 2016) or stationary phase bacteria (Pitcher *et al.*,  
269 2007b; Dupuy *et al.*, 2017) of HR-deficient *recA* mutants display IR sensitivity. This is most likely  
270 because of the involvement of *recA* in the repair of other DNA damages, in particular via its role  
271 in the SOS response, since i) spores of *B. subtilis addAB* and *recJ* mutants, deficient in the  
272 resection step of HR (no other genes coding for end resection nucleases are known to be involved  
273 in HR in *B. subtilis*), displayed IR resistance similar to the one of wild type spores (Vlašić *et al.*,  
274 2014), ii) the effects of *recA* and NHEJ mutations on IR sensitivity of *B. subtilis* and *S. ambofaciens*  
275 spores or *M. smegmatis* and *S. meliloti* stationary phase bacteria were found to be additive  
276 (Stephanou *et al.*, 2007; Pitcher *et al.*, 2007b; Moeller *et al.*, 2012; Hoff *et al.*, 2016; Dupuy *et al.*,  
277 2017). Thus, DSB repair in quiescent cells would essentially rely on NHEJ using available free non-  
278 homologous DNA ends. Furthermore, DSB repair is essential during these quiescent states  
279 because cells are more exposed to harsh environmental conditions such as desiccation, that can

280 damage the DNA. How the balance between NHEJ and HR is regulated is far from understood but  
281 crucial regarding the consequences of the repair.

282 The best characterized bacterial NHEJ regulation comes from a study aiming at characterizing  
283 gene expression in the forespore of *B. subtilis* (Wang *et al.*, 2006). The operon containing *ku<sub>Bsub</sub>*  
284 and *ligD<sub>Bsub</sub>* is expressed inside the forespore under the control of the sigma factor SigG and the  
285 transcriptional factor SpoVT. Their role in the DSB repair during spore germination (LigD is an  
286 ATP-dependent ligase and ATP is present at very low levels in spores, Korza *et al.*, 2016) has been  
287 demonstrated by the localization of a Ku-GFP fusion on the spore nucleoid after initiation of the  
288 germination, and by the high sensitivity of NHEJ deficient spores (mutated for *ku<sub>Bsub</sub>* or *ku<sub>Bsub</sub>* and  
289 *ligD<sub>Bsub</sub>*) to dry heat, a cause of DSB inside spores (Wang *et al.*, 2006). Other sources of DNA  
290 lesions, such as vacuum desiccation, heavy ions bombardment and ultraviolet radiations have a  
291 higher toxic effect on NHEJ deficient compared to wild type spores of *B. subtilis* (Moeller *et al.*,  
292 2007; Moeller *et al.*, 2008; Moeller *et al.*, 2014). According to the study of Wang *et al.*, expression  
293 of *ku<sub>Bsub</sub>* and *ligD<sub>Bsub</sub>* was observed three to four hours after *B. subtilis* cells initiate their  
294 sporulation program, concomitantly with the expression of SigG and SpoVT (Nicolas *et al.*, 2012).  
295 However, it was found that the *ku<sub>Bsub</sub> ligD<sub>Bsub</sub>* operon was slightly up-regulated (~2 times)  
296 compared to its basal level (i.e. in non sporulating conditions) when bacteria encounter a cold  
297 stress. In this condition neither SpoVT nor SigG were expressed suggesting that the NHEJ genes  
298 in *B. subtilis* could be regulated by other transcriptional factors (Nicolas *et al.*, 2012). The NHEJ  
299 proficiency of these cold stressed bacteria remains to be assessed. Besides the overexpression of  
300 the *ku<sub>Bsub</sub>* and *ligD<sub>Bsub</sub>* genes during the sporulation, lack of impact of inactivation of the HR-end  
301 resection genes on the X-ray radiation sensitivity of *B. subtilis* spores led to the proposal for a  
302 mechanism that specifically inhibits the action of AddAB and RecJ nucleases inside germinating

303 spores, rendering the NHEJ pathway the only efficient DSB repair mechanism in this condition  
304 (Vlašić *et al.*, 2014; Raguse *et al.*, 2017). Moreover, no trace of *ku* and *ligD* expression was  
305 observed at the transcriptional (Nicolas *et al.*, 2012) and translational (Sinai *et al.*, 2015) levels  
306 during the revival of spores. Ku (Ku-GFP fusion) expressed during sporulation is detectable  
307 associated to the nucleoid during the first 30 min of the spore revival (Wang *et al.*, 2006), i.e.  
308 before initiation of DNA replication and expression of DNA repair genes such as *recA* (Sinai *et al.*,  
309 2015). Then, *in B. subtilis*, NHEJ and HR do not seem to compete for the repair of two dsDNA  
310 ends.

311 NHEJ is not a hallmark of bacterial sporulation since *ku* and *ligD* genes are also present in non-  
312 spore forming bacteria and all spore-forming bacteria do not encode these proteins (Aravind and  
313 Koonin, 2001; Bowater and Doherty, 2006; McGovern *et al.*, 2016). In Mycobacteria and in *S.*  
314 *meliloti*, *ku* and *ligD* inactivation leads to a sensitization against IR of mutant bacteria grown in  
315 stationary phase and have no or little effect in logarithmic phase, even if *recA* is inactivated  
316 (Stephanou *et al.*, 2007; Pitcher *et al.*, 2007b; Kobayashi *et al.*, 2008; Dupuy *et al.*, 2017). On the  
317 other hand, *ku* and/or *ligD* inactivation strongly reduces the capacity of logarithmically grown  
318 Mycobacteria or *S. meliloti* to repair enzymatically linearized plasmid DNA upon transformation  
319 (Gong *et al.*, 2004; Gong *et al.*, 2005; Aniukwu *et al.*, 2008; Dupuy *et al.*, 2018) as well as  
320 enzymatically induced DSB on the chromosome (Stephanou *et al.*, 2007; Gupta *et al.*, 2011;  
321 Dupuy *et al.*, 2018). Moreover, *ligD* is required in exponentially growing *M. smegmatis* bacteria  
322 for infection by the mycobacteriophages Omega and Corndog, where it is thought to promote  
323 circularization at the *cos* site of the phage genomes, thanks to the Ku expressed by the phages  
324 (Pitcher *et al.*, 2006). Altogether, these observations show that although not involved in the  
325 repair of DSB induced by IR in logarithmic phase, Ku- and LigD-dependent NHEJ is well active in



326 this growth stage in several bacterial species. This is consistent with the detection of similar  
327 amounts of Ku and similar activity of its promoter in logarithmic and early stationary phases in  
328 *M. smegmatis* (Bhattarai *et al.*, 2014; Zhou *et al.*, 2015), even if an increased activity of the Ku  
329 promoter was observed in late stationary phase cells (Bhattarai *et al.*, 2014). This is also  
330 consistent with the expression of a *S. meliloti ku2-lacZ* transcriptional fusion in both log and  
331 stationary phase, and the visualization of Ku2-YFP foci after X-ray treatment of cells in both  
332 phases (Kobayashi *et al.*, 2008). Together, these observations suggest that the action of NHEJ for  
333 repair of IR induced DSBs, which have most probably similar molecular configurations in  
334 logarithmic and stationary phases, is negatively regulated in logarithmic phase. The mechanism  
335 of this inhibition does not seem dependent on RecA since the deletion of *recA* in *M. smegmatis*  
336 does not increase the NHEJ (Gupta *et al.*, 2011), however it could involve the competition with  
337 presynaptic actors of the HR.

338 In *S. meliloti*, NHEJ activity as measured using a linear plasmid repair assay, displays a significant  
339 stimulation when bacteria enter stationary phase, or are exposed to a thermic stress (Dupuy *et al.*,  
340 2018). This stimulation mainly results from the upregulation of *ku* gene expression. The *ku3*  
341 *ku4* operon, as well as the *ligD4* gene, appeared to be under transcriptional control of the general  
342 stress response EcfG sigma factor RpoE2. Examination of *ku* promoter regions in  $\alpha$ -  
343 Proteobacteria revealed the presence of putative EcfG binding sites upstream at least one *ku*  
344 gene in about 90 % of the species, suggesting that stress control of *ku* and *ligD* transcription is  
345 widespread in this bacterial class. Accordingly, stress induction of *ku* and/or *ligD* gene expression  
346 was described in *Mesorhizobium loti* and *Bradyrhizobium japonicum* (Cytryn *et al.*, 2007;  
347 Alexandre *et al.*, 2014). How *ku2* is up-regulated in stress conditions is currently unknown, but it  
348 appears not to be part of the general stress response, the heat shock response, the SOS response,

349 the stringent response nor to be regulated by quorum sensing (Kobayashi *et al.*, 2008; Dupuy *et*  
350 *al.*, 2018). Whether stress control of NHEJ gene expression is conserved beyond  $\alpha$ -Proteobacteria  
351 remains to be tested experimentally. In *Streptomyces*, major NHEJ genes are expressed in late-  
352 growth phase (Hoff *et al.*, 2016) and gene products accumulate in spores. However, no induction  
353 of the NHEJ-like genes could be detected in *S. ambofaciens* in response to DNA damaging agents  
354 (i.e., Mitomycin C, Hoff *et al.*, 2016).

355 Besides transcriptional regulation, post-translational modification of NHEJ proteins could also  
356 regulate NHEJ activity in bacteria. Indeed the deacetylase Sir2 from *M. smegmatis* is required for  
357 an efficient NHEJ repair *in vivo* (Li *et al.*, 2011). Overexpression of *sir2* leads to a decreased  
358 amount of Ku in bacteria, without altering the quantity of the corresponding mRNA. In line with  
359 this interaction, Ku<sub>*Msm*</sub> is acetylated at least at the residue K29, which is conserved among  
360 bacterial Ku, and this modification is mainly observed in stationary phase (Zhou *et al.*, 2015). Even  
361 if the mechanism by which Sir2 regulates NHEJ proteins remains to be clarified, it seems likely  
362 that acetylation could regulate the activity of NHEJ (Zhou *et al.*, 2015).

#### 363 **IV Cross talk with other DNA repair mechanisms**

364 As mentioned in the section III, DSB repair in quiescent cells would essentially rely on NHEJ, but  
365 NHEJ and HR repair mechanisms may both co-exist, at least in some NHEJ proficient bacteria  
366 during replicating phases. Moreover, in *Streptomyces*, recent studies (Hoff *et al.*, 2018) on the  
367 consequences of DSB repair in the terminal region of the linear chromosome revealed events of  
368 HR repair involving homologous sequences at distinct loci suggesting that HR can occur even in  
369 the absence of a replicated copy of the chromosome (quiescent phases). This raises the question  
370 of how the choice between the two pathways is performed and if there is a competition between

371 both mechanisms. In eukaryotes, it is known that Ku proteins rapidly sequester DSB free ends  
372 and shield them from unscheduled exonucleases, preventing HR to proceed. The removal of Ku  
373 involves multiple pathways including ubiquitination, neddylation, phosphorylation and the  
374 nuclease activity of the MRE11-dependent complex MRX (yeast) /MRN (mammals) which plays a  
375 pivotal role in NHEJ/HR balance (for review see Shibata, 2017). Similarly, in *Mycobacterium*, *in*  
376 *vitro* analyses revealed that Ku<sub>Mtub</sub> protects DSBs from resection by AdnAB helicase-nuclease  
377 (Sinha *et al.*, 2009) and Ku<sub>Mtub</sub> but not LigD<sub>Mtub</sub> inactivation is compensated by an elevated HR *in*  
378 *vivo* (Gupta *et al.*, 2011). Likewise, in *S. avermitilis*, the deletion of *kuSave1* and *kuSave2* seemed  
379 associated with an increased efficiency of HR (Zhang *et al.*, 2012). Recent electronic beam  
380 irradiation assays in *S. ambofaciens* (Hoff *et al.*, 2016, Leblond *et al.*, unpublished) also suggest  
381 that NHEJ inactivation (*kuSamb* triple deletion or *ligCSamb* and *ligDSamb* double deletion) is  
382 compensated for by an increased HR repair. However, unlike in *Mycobacterium* where loss of HR  
383 is not compensated by increased NHEJ, a reciprocal effect is observed in *S. ambofaciens* since a  
384 *recA* mutant is 100 times more affected by NHEJ inactivation than the wild-type strain.

385 Ku and LigD proteins may help coupling NHEJ to other DNA repair pathways. IR-induced cellular  
386 damages on DNA include oxidized bases, abasic sites and single-strand breaks (SSB). However,  
387 due to the presumed “clustered damage” induced by IR (Ward, 1994), several DNA lesions can  
388 be present in close proximity and, if they are not repaired efficiently, can lead to DSB (Ward,  
389 1994; Blaisdell and Wallace, 2001; Cox and Battista, 2005). Then, DSBs induced by IR are  
390 potentially surrounded by multiple other DNA lesions. Part of these lesions are taken in charge  
391 by the base excision repair (BER) pathway, the nicking of the phosphodiester bond 3’ to an  
392 apurinic/apyrimidinic (AP) site by an AP-lyase being one possible step of this pathway (for review  
393 see Almeida and Sobol, 2007). Human Ku, besides its function as a scaffolder of the NHEJ

394 machinery, is also an AP-lyase suggested to be involved in the repair of AP site near DSBs (Roberts  
395 *et al.*, 2010). This activity is conserved in Ku from *B. subtilis* and *P. aeruginosa* (de Ory *et al.*,  
396 2014). As Ku<sub>Bsub</sub> is able to thread inward DNA molecules (McGovern *et al.*, 2016), via loading on  
397 their broken ends, this could facilitate the repair of abasic sites surrounding a DSB. The interplay  
398 between NHEJ and BER in bacteria is reinforced by the discovery of an AP-lyase activity driven by  
399 the other NHEJ factor: LigD (de Ory *et al.*, 2016). The involvement of LigC1<sub>Msm</sub> in BER during  
400 stationary phase of *M. smegmatis* (Płociński *et al.*, 2017) shows that the link between NHEJ and  
401 BER can even be extended to alternative NHEJ proteins. This multiplicity of NHEJ actors involved  
402 in BER is not specific to bacteria, human DNA polymerase  $\lambda$  is one example (see Bebenek *et al.*,  
403 2014).

404 Remarkably, Ku<sub>Bsub</sub> is also able to linearize a supercoiled plasmid containing AP-sites (de Ory *et*  
405 *al.*, 2014). This counter-intuitive effect for a protein intended to repair DSB is supposed to be  
406 dependent on its AP-lyase activity and its ability to bind end-free dsDNA molecules, DSB arising  
407 by the nicking of phosphodiester bonds of proximal AP-sites on opposite strands of a dsDNA  
408 molecules containing numerous such lesions (de Ory *et al.*, 2014). Then it is tempting to propose  
409 that NHEJ could have evolved from a Ku-dependent BER pathway, primarily independent on the  
410 presence of DSB, that has acquired the ability to join Ku-catalyzed broken DNA ends, via the  
411 recruitment of LigD, in conditions where the bacterial genome contains numerous AP-sites. This  
412 hypothesis and the fact that most of the bacterial Ku proteins contain an extended C-ter tail that  
413 is required for their interaction on end-free DNA (Kushwaha and Grove, 2013b; McGovern *et al.*,  
414 2016) but for which the deletion has only a mild effect *in vitro* on the NHEJ efficiency (Zhu and  
415 Shuman, 2010; McGovern *et al.*, 2016) suggest that the extended C-ter of Ku could play an  
416 important role in a Ku-dependent BER acting on AP-sites embedded in duplex regions of DNA.

417 A recent study has functionally linked the NHEJ of *P. aeruginosa* to the MutL protein, a  
418 component of the mismatch repair (MMR) pathway. Searching for the molecular mechanism at  
419 the basis of a large genomic deletion in the *P. aeruginosa* chromosome allowing to prevent  
420 infection by a bacteriophage, Shen *et al.* (2018) have shown that MutL<sub>Paer</sub> is required for getting  
421 the deletion. MutL<sub>Paer</sub> is an endonuclease promoting SSB and, directly or indirectly, DSB (as a  
422 consequence of multiple SSB in dsDNA). Surprisingly, the MutL<sub>Paer</sub>-dependent genomic deletion  
423 observed after the bacteriophage infection is not dependent on MutS<sub>Paer</sub>, the partner of MutL<sub>Paer</sub>  
424 in MMR, suggesting a new function of MutL in this bacterium. Whatever the mechanism, the  
425 reduced frequency of the corresponding deletion in a NHEJ deficient mutant led to the proposal  
426 that DSB generated by MutL<sub>Paer</sub> are repaired by the NHEJ pathway in this bacterium (Shen *et al.*,  
427 2018). On the other hand, the majority of archaeal species and many actinobacteria lack MutHLS  
428 system but display genes putatively involved in a novel mismatch repair system, called  
429 NucS/EndoMS, which has been shown in *Corynebacterium glutamicum* to cleave both strands at  
430 double-strand DNA mismatched substrate during the repair process (Castañeda-García *et al.*,  
431 2017; Ishino *et al.*, 2018). Whether these induced DSBs are thereafter processed by NHEJ or HR  
432 is not yet known, but one can suggest that chromosomes of bacteria possessing a NucS/EndoMS  
433 system are more prone to DSBs and consequently to mutations and rearrangements, especially  
434 if they also feature a NHEJ machinery, like many actinomycetes.

435 Another mechanism that induces DSBs is the type II CRISPR-Cas (Clustered Regularly Interspaced  
436 Short Palindromic Repeats and associated proteins) system which is conferring adaptive  
437 immunity against phages and exogenous mobile genetic elements (for review see Marraffini,  
438 2015) and which has been engineered to provide an efficient genome editing method (Jinek *et*  
439 *al.*, 2012). Since the CRISPR-Cas system generates DSBs, it could display antagonistic interactions

440 with DSB repair systems. Bernheim and co-authors (2017) revealed that NHEJ and type II-A  
441 CRISPR-Cas systems do not co-occur in all but one of more than 5000 sequenced bacterial  
442 genomes. On the other hand, this strong avoidance of co-occurrence was not observed with type  
443 II-C system, which is consistent with the fact that, unlike type II-A system, type II-C system lacks  
444 Csn2, a multimeric toroidal binding protein that binds double stranded DNA ends and has similar  
445 properties to Ku protein (Arslan *et al.*, 2013). The binding of Csn2 at DNA ends could block access  
446 to Ku and inhibit efficient repair by NHEJ. In contrast, *B. subtilis* NHEJ system does not impair  
447 *Streptococcus pyogenes* or *Streptococcus thermophilus* type II-A CRISPR-Cas9 system (Bernheim  
448 *et al.*, 2017).

#### 449 **V- NHEJ and genome evolution**

450 NHEJ is inherently unfaithful, as DNA ends of DSB are often remodeled before ligation, thus  
451 generating short insertions or deletions at the repair junction (although repair fidelity seems to  
452 vary depending on the tested organism, the type of DNA ends being repaired or the enzymes  
453 involved). Interestingly, NHEJ is generally more active in stationary phase or during sporulation,  
454 and in  $\alpha$ -proteobacteria, is activated under stress conditions, in a process dependent on the  
455 general stress response sigma factor EcfG. This is reminiscent of the so-called stress-induced  
456 mutagenesis of enterobacteria, where the highly faithful HR-dependent DSB repair becomes  
457 mutagenic in a process dependent on the general stress response sigma factor RpoS (Shee *et al.*,  
458 2011). NHEJ was also recently suggested to participate in stationary phase mutagenesis in  
459 *Pseudomonas putida* (Paris *et al.*, 2015). NHEJ could then significantly contribute to DSB-  
460 dependent stress-induced mutagenesis in bacteria and could thus participate in their adaptation  
461 to adverse environmental conditions.

462 In addition to the short, non-templated nucleotide insertions usually observed at NHEJ repair  
463 junctions, insertion of longer DNA sequences (several tens of nucleotides) of either plasmid or  
464 chromosomal origin, was described to have occurred in rare instances at plasmid or chromosome  
465 DSB repair junctions in *M. smegmatis* and *S. ambofaciens* (Akey *et al.*, 2006; Aniukwu *et al.*, 2008;  
466 Hoff *et al.*, 2018). Although the mechanism of such DNA insertions was not clearly established, it  
467 may have resulted from the incidental capture of indigenous DNA fragments at the break by the  
468 NHEJ machinery. In eukaryotes, DNA inserted into the genome by NHEJ during evolution include  
469 foreign DNA fragments such as mitochondrial DNA, transposable elements, and viral DNA (Moore  
470 and Haber, 1996; Ricchetti *et al.*, 1999; Lin and Waldman, 2001a; Lin and Waldman, 2001b; Nakai  
471 *et al.*, 2003; Hazkani-Covo and Covo, 2008). In yeast, capture of extranuclear DNA at repaired  
472 DSBs was shown to be dependent on DNA end resection and NHEJ complexes (Decottignies,  
473 2005). In the same way, the low level of genome rearrangements in the yeast *Lachancea kluyveri*  
474 since its separation from its ancestor with *S. cerevisiae* was assigned to the loss of NHEJ genes in  
475 this phylum (Gordon *et al.*, 2011). Interestingly, Dupuy *et al.* demonstrated that linear DNA of  
476 heterologous origin, introduced in *S. meliloti* by transformation, could be effectively captured in  
477 a NHEJ-dependent way at an artificially generated DSB in the chromosome (Dupuy *et al.*, 2018).  
478 In *Streptomyces*, since the variable NHEJ genes encode actors of the response to DNA damage  
479 (Hoff *et al.*, 2016) that presumably participate to the NHEJ and other DNA repair pathways, it is  
480 likely that different combinations of NHEJ genes result in variable recombination profiles, and  
481 consequently in variable abilities to integrate DNA from horizontal gene transfer. Like for *E. coli*  
482 mutator strains (Giraud *et al.*, 2001), such recombination proficient lineages could provide a  
483 flexible adaption property to the population in the ecological context by allowing rapid and  
484 intense diversification of the accessory gene set. This speculation is substantiated by the strong  
485 variability of the NHEJ-gene set revealed within a *Streptomyces* population originating from a

486 forest soil (Leblond *et al.*, unpublished). Altogether, these data suggest that NHEJ could  
487 participate in the acquisition of heterologous DNA during lateral transfer events, overcoming the  
488 lack of sequence homology between distantly related species. Interestingly, in support of this  
489 hypothesis, Popa *et al.* (2011) found that gene transfer events from distant donors are more  
490 frequent when recipient genomes encode NHEJ proteins. In addition, Kidane *et al.* (2009) have  
491 observed that the absence of Ku reduces natural plasmid transformation and phage DNA  
492 transfection in *B. subtilis*.

493

#### 494 **Concluding remarks**

495 As mentioned within this review, many NHEJ proficient bacteria contain several stand-alone  
496 genes displaying similarities with genes coding for NucDom, Poldom, LigDom and/or *ku* genes  
497 and their multiplicity renders difficult the identification of the *bona fide* NHEJ genes. The use of  
498 two-hybrid screening coupled with affinity purification of the products of these genes in their  
499 host and in conditions where NHEJ is active should help to characterize actual NHEJ complexes,  
500 and potentially identify new components of the NHEJ machinery, or new links with other DNA  
501 repair pathways (e.g. role of AP-lyase activities in response to genotoxic stresses). Structural  
502 information (i.e. cryo-electron microscopy and super-resolution microscopy analyses) associated  
503 to 3D structures of stand-alone NHEJ actors (LigD, Ku) and of the complete bacterial NHEJ  
504 machinery will be of crucial importance for our basic understanding of this mechanism (e.g. NHEJ  
505 synapsis step, DNA extremities tethering, processing of variable DNA ends). Furthermore,  
506 alternative NHEJ machineries may be involved depending on the growth phase or growth  
507 conditions (see above the stress-induced NHEJ machineries of *S. meliloti*). This raises the  
508 questions of the competition between distinct NHEJ machineries, and between HR and NHEJ, to



509 repair DSBs in conditions where both mechanisms coexist. It is clear, based on our knowledge of  
510 the eukaryotic NHEJ, that DNA ends accessibility should be a key point that affects the HR/NHEJ  
511 balance. Genetic, biochemical and cellular localization studies of proteins involved in early steps  
512 of these pathways over time after DSB induction should help to elucidate preferred routes of DSB  
513 repair in different life states. Further, experimental evolution could constitute a powerful  
514 approach for testing the involvement of NHEJ in integration of DNA acquired by horizontal gene  
515 transfer and its impact on the capacity to create and maintain genetic diversity within bacterial  
516 populations in their ecosystems.

517 Finally, understanding the bacterial NHEJ may have therapeutic interest. Indeed, many human  
518 pathogens possess a NHEJ repair system. This is for instance the case for the most deadly human  
519 pathogen *M. tuberculosis* and the nosocomial pathogen *P. aeruginosa*. Although NHEJ is not  
520 directly involved in survival or virulence in different mammal hosts (Brzostek *et al.*, 2014; Heaton  
521 *et al.*, 2014), it could provide a way to survive to genotoxic conditions in certain non-dividing  
522 states. Identifying these states and developing anti-NHEJ compounds based on the high-  
523 resolution structure of a NHEJ machinery could help to fight against these pathogens.

524

525

526 **Acknowledgments**

527 This work was supported by INRA (Institut National de la Recherche Agronomique) -including by  
528 the Département Santé des Plantes et Environnement-, by the French Laboratory of Excellence  
529 project TULIP (ANR-10-LABX-4), by the French National Research Agency program Streptoflux  
530 (ANR Blanc 0096\_01), and by ANR through the Laboratory of Excellence ARBRE (ANR-12-  
531 LABXARBRE-01). The authors declare that they have no competing interests.

532

533 **References**

534 Akey, D., Martins, A., Aniuoku, J., Glickman, M.S., Shuman, S., and Berger, J.M. (2006) Crystal  
535 structure and nonhomologous end-joining function of the ligase component of Mycobacterium DNA  
536 ligase D. *J Biol Chem* **281**: 13412–13423.

537 Alexandre, A., Laranjo, M., and Oliveira, S. (2014) Global Transcriptional Response to Heat Shock of  
538 the Legume Symbiont Mesorhizobium loti MAFF303099 Comprises Extensive Gene Downregulation.  
539 *DNA Res* **21**: 195–206.

540 Almeida, K.H., and Sobol, R.W. (2007) A unified view of base excision repair: lesion-dependent  
541 protein complexes regulated by post-translational modification. *DNA Repair (Amst)* **6**: 695–711.

542 Aniuoku, J., Glickman, M.S., and Shuman, S. (2008) The pathways and outcomes of mycobacterial  
543 NHEJ depend on the structure of the broken DNA ends. *Genes Dev* **22**: 512–527.

544 Aravind, L., and Koonin, E.V. (2001) Prokaryotic homologs of the eukaryotic DNA-end-binding protein  
545 Ku, novel domains in the Ku protein and prediction of a prokaryotic double-strand break repair  
546 system. *Genome Res* **11**: 1365–1374.

547 Arslan, Z., Wurm, R., Brener, O., Ellinger, P., Nagel-Steger, L., Oesterhelt, F., *et al.* (2013) Double-  
548 strand DNA end-binding and sliding of the toroidal CRISPR-associated protein Csn2. *Nucleic Acids Res*  
549 **41**: 6347–6359.

550 Aylon, Y., Liefshitz, B., and Kupiec, M. (2004) The CDK regulates repair of double-strand breaks by  
551 homologous recombination during the cell cycle. *EMBO J* **23**: 4868–4875.

552 Bartlett, E.J., Brissett, N.C., and Doherty, A.J. (2013) Ribonucleolytic resection is required for repair of  
553 strand displaced nonhomologous end-joining intermediates. *Proc Natl Acad Sci USA* **110**: E1984-  
554 1991.

555 Bartlett, E.J., Brissett, N.C., Plocinski, P., Carlberg, T., and Doherty, A.J. (2016) Molecular basis for  
556 DNA strand displacement by NHEJ repair polymerases. *Nucleic Acids Res* **44**: 2173–2186.

557 Bebenek, K., Pedersen, L.C., and Kunkel, T.A. (2014) Structure-function studies of DNA polymerase  $\lambda$ .  
558 *Biochemistry* **53**: 2781–2792.

559 Bernheim, A., Calvo-Villamañán, A., Basier, C., Cui, L., Rocha, E.P.C., Touchon, M., and Bikard, D.  
560 (2017) Inhibition of NHEJ repair by type II-A CRISPR-Cas systems in bacteria. *Nat Commun* **8**: 2094.

561 Bhattacharyya, S., Soniat, M.M., Walker, D., Jang, S., Finkelstein, I.J., and Harshey, R.M. (2018) Phage  
562 Mu Gam protein promotes NHEJ in concert with Escherichia coli ligase. *Proc Natl Acad Sci USA* **115**:  
563 E11614–E11622.

564 Bhattarai, H., Gupta, R., and Glickman, M.S. (2014) DNA ligase C1 mediates the LigD-independent  
565 nonhomologous end-joining pathway of Mycobacterium smegmatis. *J Bacteriol* **196**: 3366–3376.

566 Blaisdell, J.O., and Wallace, S.S. (2001) Abortive base-excision repair of radiation-induced clustered  
567 DNA lesions in Escherichia coli. *Proc Natl Acad Sci USA* **98**: 7426–7430.

568 Bowater, R., and Doherty, A.J. (2006) Making Ends Meet: Repairing Breaks in Bacterial DNA by Non-  
569 Homologous End-Joining. *PLoS Genet* **2** <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC1378119/>.

570 Brissett, N.C., and Doherty, A.J. (2009) Repairing DNA double-strand breaks by the prokaryotic non-  
571 homologous end-joining pathway. *Biochem Soc Trans* **37**: 539–545.

572 Brissett, N.C., Martin, M.J., Bartlett, E.J., Bianchi, J., Blanco, L., and Doherty, A.J. (2013) Molecular  
573 basis for DNA double-strand break annealing and primer extension by an NHEJ DNA polymerase. *Cell*  
574 *Rep* **5**: 1108–1120.

575 Brissett, N.C., Martin, M.J., Pitcher, R.S., Bianchi, J., Juarez, R., Green, A.J., *et al.* (2011) Structure of a  
576 preternary complex involving a prokaryotic NHEJ DNA polymerase. *Mol Cell* **41**: 221–231.

577 Brzostek, A., Szulc, I., Klink, M., Brzezinska, M., Sulowska, Z., and Dziadek, J. (2014) Either non-  
578 homologous ends joining or homologous recombination is required to repair double-strand breaks in  
579 the genome of macrophage-internalized Mycobacterium tuberculosis. *PLoS ONE* **9**: e92799.

580 Castañeda-García, A., Prieto, A.I., Rodríguez-Beltrán, J., Alonso, N., Cantillon, D., Costas, C., *et al.*  
581 (2017) A non-canonical mismatch repair pathway in prokaryotes. *Nat Commun* **8**: 14246.

582 Chang, H.H.Y., Pannunzio, N.R., Adachi, N., and Lieber, M.R. (2017) Non-homologous DNA end joining  
583 and alternative pathways to double-strand break repair. *Nat Rev Mol Cell Biol* **18**: 495–506.

584 Chayot, R., Montagne, B., Mazel, D., and Ricchetti, M. (2010) An end-joining repair mechanism in  
585 Escherichia coli. *Proc Natl Acad Sci USA* **107**: 2141–2146.

586 Clerici, M., Mantiero, D., Guerini, I., Lucchini, G., and Longhese, M.P. (2008) The Yku70–Yku80  
587 complex contributes to regulate double-strand break processing and checkpoint activation during the  
588 cell cycle. *EMBO Rep* **9**: 810–818.

589 Cox, M.M., and Battista, J.R. (2005) Deinococcus radiodurans - the consummate survivor. *Nat Rev*  
590 *Microbiol* **3**: 882–892.

591 Cytryn, E.J., Sangurdekar, D.P., Streeter, J.G., Franck, W.L., Chang, W., Stacey, G., *et al.* (2007)  
592 Transcriptional and Physiological Responses of Bradyrhizobium japonicum to Desiccation-Induced  
593 Stress. *J Bacteriol* **189**: 6751–6762.

594 Decottignies, A. (2005) Capture of extranuclear DNA at fission yeast double-strand breaks. *Genetics*  
595 **171**: 1535–1548.

596 Della, M., Palmboos, P.L., Tseng, H.-M., Tonkin, L.M., Daley, J.M., Topper, L.M., *et al.* (2004)  
597 Mycobacterial Ku and ligase proteins constitute a two-component NHEJ repair machine. *Science* **306**:  
598 683–685.

599 Doherty, A.J., Jackson, S.P., and Weller, G.R. (2001) Identification of bacterial homologues of the Ku  
600 DNA repair proteins. *FEBS Lett* **500**: 186–188.

601 Downs, J.A., and Jackson, S.P. (2004) A means to a DNA end: the many roles of Ku. *Nat Rev Mol Cell*  
602 *Biol* **5**: 367–378.

603 Dupuy, P., Gourion, B., Sauviac, L., and Bruand, C. (2017) DNA double-strand break repair is involved  
604 in desiccation resistance of *Sinorhizobium meliloti*, but is not essential for its symbiotic interaction  
605 with *Medicago truncatula*. *Microbiology (Reading, Engl)* **163**: 333–342.

606 Dupuy, P., Sauviac, L., and Bruand, C. (2018) Stress-inducible NHEJ in bacteria: function in DNA repair  
607 and acquisition of heterologous DNA. *Nucleic Acids Res* .

608 Fagagna, F. d'Adda di, Weller, G.R., Doherty, A.J., and Jackson, S.P. (2003) The Gam protein of  
609 bacteriophage Mu is an orthologue of eukaryotic Ku. *EMBO Rep* **4**: 47–52.

610 Galibert, F., Finan, T.M., Long, S.R., Puhler, A., Abola, P., Ampe, F., *et al.* (2001) The composite  
611 genome of the legume symbiont *Sinorhizobium meliloti*. *Science* **293**: 668–672.

612 Giraud, A., Matic, I., Tenailon, O., Clara, A., Radman, M., Fons, M., and Taddei, F. (2001) Costs and  
613 benefits of high mutation rates: adaptive evolution of bacteria in the mouse gut. *Science* **291**: 2606–  
614 2608.

615 Gong, C., Bongiorno, P., Martins, A., Stephanou, N.C., Zhu, H., Shuman, S., and Glickman, M.S. (2005)  
616 Mechanism of nonhomologous end-joining in mycobacteria: a low-fidelity repair system driven by  
617 Ku, ligase D and ligase C. *Nat Struct Mol Biol* **12**: 304–312.

618 Gong, C., Martins, A., Bongiorno, P., Glickman, M., and Shuman, S. (2004) Biochemical and genetic  
619 analysis of the four DNA ligases of mycobacteria. *J Biol Chem* **279**: 20594–20606.

620 Gordon, J.L., Byrne, K.P., and Wolfe, K.H. (2011) Mechanisms of chromosome number evolution in  
621 yeast. *PLoS Genet* **7**: e1002190.

622 Gu, J., and Lieber, M.R. (2008) Mechanistic flexibility as a conserved theme across 3 billion years of  
623 nonhomologous DNA end-joining. *Genes Dev* **22**: 411–415.

624 Gupta, R., Barkan, D., Redelman-Sidi, G., Shuman, S., and Glickman, M.S. (2011) Mycobacteria exploit  
625 three genetically distinct DNA double-strand break repair pathways. *Mol Microbiol* **79**  
626 <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC3812669/>.

627 Hazkani-Covo, E., and Covo, S. (2008) Numt-mediated double-strand break repair mitigates deletions  
628 during primate genome evolution. *PLoS Genet* **4**: e1000237.

629 Heaton, B.E., Barkan, D., Bongiorno, P., Karakousis, P.C., and Glickman, M.S. (2014) Deficiency of  
630 double-strand DNA break repair does not impair *Mycobacterium tuberculosis* virulence in multiple  
631 animal models of infection. *Infect Immun* **82**: 3177–3185.

632 Hoff, G., Bertrand, C., Piotrowski, E., Thibessard, A., and Leblond, P. (2018) Genome plasticity is  
633 governed by double strand break DNA repair in *Streptomyces*. *Sci Rep* **8**: 5272.

- 634 Hoff, G., Bertrand, C., Zhang, L., Piotrowski, E., Chipot, L., Bontemps, C., *et al.* (2016) Multiple and  
635 Variable NHEJ-Like Genes Are Involved in Resistance to DNA Damage in *Streptomyces ambofaciens*.  
636 *Front Microbiol* **7**: 1901.
- 637 Ira, G., Pellicoli, A., Balijja, A., Wang, X., Fiorani, S., Carotenuto, W., *et al.* (2004) DNA end resection,  
638 homologous recombination and DNA damage checkpoint activation require CDK1. *Nature* **431**: 1011–  
639 1017.
- 640 Ishino, S., Skouloubris, S., Kudo, H., Stead, C. l’Hermitte-, Es-Sadik, A., Lambry, J.-C., *et al.* (2018)  
641 Activation of the mismatch-specific endonuclease EndoMS/NucS by the replication clamp is required  
642 for high fidelity DNA replication. *Nucleic Acids Res* **46**: 6206–6217.
- 643 Jinek, M., Chylinski, K., Fonfara, I., Hauer, M., Doudna, J.A., and Charpentier, E. (2012) A  
644 Programmable Dual-RNA–Guided DNA Endonuclease in Adaptive Bacterial Immunity. *Science* **337**:  
645 816–821.
- 646 Kidane, D., Carrasco, B., Manfredi, C., Rothmaier, K., Ayora, S., Tadesse, S., *et al.* (2009) Evidence for  
647 different pathways during horizontal gene transfer in competent *Bacillus subtilis* cells. *PLoS Genet* **5**:  
648 e1000630.
- 649 Kobayashi, H., Simmons, L.A., Yuan, D.S., Broughton, W.J., and Walker, G.C. (2008) Multiple Ku  
650 orthologues mediate DNA non-homologous end-joining in the free-living form and during chronic  
651 infection of *Sinorhizobium meliloti*. *Mol Microbiol* **67**: 350–363.
- 652 Korza, G., Setlow, B., Rao, L., Li, Q., and Setlow, P. (2016) Changes in *Bacillus* Spore Small Molecules,  
653 rRNA, Germination, and Outgrowth after Extended Sublethal Exposure to Various Temperatures:  
654 Evidence that Protein Synthesis Is Not Essential for Spore Germination. *J Bacteriol* **198**: 3254–3264.
- 655 Kushwaha, A.K., and Grove, A. (2013a) C-terminal low-complexity sequence repeats of  
656 *Mycobacterium smegmatis* Ku modulate DNA binding. *Biosci Rep* **33**: 175–184.
- 657 Kushwaha, A.K., and Grove, A. (2013b) *Mycobacterium smegmatis* Ku binds DNA without free ends.  
658 *Biochem J* **456**: 275–282.
- 659 Li, Z., Wen, J., Lin, Y., Wang, S., Xue, P., Zhang, Z., *et al.* (2011) A Sir2-Like Protein Participates in  
660 *Mycobacterial* NHEJ. *PLoS One* **6** <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC3102665/>.
- 661 Lin, Y., and Waldman, A.S. (2001a) Capture of DNA sequences at double-strand breaks in mammalian  
662 chromosomes. *Genetics* **158**: 1665–1674.
- 663 Lin, Y., and Waldman, A.S. (2001b) Promiscuous patching of broken chromosomes in mammalian  
664 cells with extrachromosomal DNA. *Nucleic Acids Res* **29**: 3975–3981.
- 665 Malyarchuk, S., Wright, D., Castore, R., Klepper, E., Weiss, B., Doherty, A.J., and Harrison, L. (2007)  
666 Expression of *Mycobacterium tuberculosis* Ku and Ligase D in *Escherichia coli* results in RecA and  
667 RecB-independent DNA end-joining at regions of microhomology. *DNA Repair (Amst)* **6**: 1413–1424.
- 668 Marraffini, L.A. (2015) CRISPR-Cas immunity in prokaryotes. *Nature* **526**: 55–61.
- 669 McGovern, S., Baconnais, S., Roblin, P., Nicolas, P., Drevet, P., Simonson, H., *et al.* (2016) C-terminal  
670 region of bacterial Ku controls DNA bridging, DNA threading and recruitment of DNA ligase D for  
671 double strand breaks repair. *Nucleic Acids Res* **44**: 4785–4806.

672 Moeller, R., Raguse, M., Reitz, G., Okayasu, R., Li, Z., Klein, S., *et al.* (2014) Resistance of *Bacillus*  
673 *subtilis* spore DNA to lethal ionizing radiation damage relies primarily on spore core components and  
674 DNA repair, with minor effects of oxygen radical detoxification. *Appl Environ Microbiol* **80**: 104–109.

675 Moeller, R., Reitz, G., Li, Z., Klein, S., and Nicholson, W.L. (2012) Multifactorial resistance of *Bacillus*  
676 *subtilis* spores to high-energy proton radiation: role of spore structural components and the  
677 homologous recombination and non-homologous end joining DNA repair pathways. *Astrobiology* **12**:  
678 1069–1077.

679 Moeller, R., Setlow, P., Horneck, G., Berger, T., Reitz, G., Rettberg, P., *et al.* (2008) Roles of the Major,  
680 Small, Acid-Soluble Spore Proteins and Spore-Specific and Universal DNA Repair Mechanisms in  
681 Resistance of *Bacillus subtilis* Spores to Ionizing Radiation from X Rays and High-Energy Charged-  
682 Particle Bombardment. *J Bacteriol* **190**: 1134–1140.

683 Moeller, R., Stackebrandt, E., Reitz, G., Berger, T., Rettberg, P., Doherty, A.J., *et al.* (2007) Role of DNA  
684 Repair by Nonhomologous-End Joining in *Bacillus subtilis* Spore Resistance to Extreme Dryness,  
685 Mono- and Polychromatic UV, and Ionizing Radiation. *J Bacteriol* **189**: 3306–3311.

686 Moore, J.K., and Haber, J.E. (1996) Capture of retrotransposon DNA at the sites of chromosomal  
687 double-strand breaks. *Nature* **383**: 644–646.

688 Nair, P.A., Smith, P., and Shuman, S. (2010) Structure of bacterial LigD 3'-phosphoesterase unveils a  
689 DNA repair superfamily. *Proc Natl Acad Sci USA* **107**: 12822–12827.

690 Nakai, H., Montini, E., Fuess, S., Storm, T.A., Grompe, M., and Kay, M.A. (2003) AAV serotype 2  
691 vectors preferentially integrate into active genes in mice. *Nat Genet* **34**: 297–302.

692 Nicolas, P., Mäder, U., Dervyn, E., Rochat, T., Leduc, A., Pigeonneau, N., *et al.* (2012) Condition-  
693 dependent transcriptome reveals high-level regulatory architecture in *Bacillus subtilis*. *Science* **335**:  
694 1103–1106.

695 Ory, A. de, Nagler, K., Carrasco, B., Raguse, M., Zafra, O., Moeller, R., and Vega, M. de (2016)  
696 Identification of a conserved 5'-dRP lyase activity in bacterial DNA repair ligase D and its potential  
697 role in base excision repair. *Nucleic Acids Res* **44**: 1833–1844.

698 Ory, A. de, Zafra, O., and Vega, M. de (2014) Efficient processing of abasic sites by bacterial  
699 nonhomologous end-joining Ku proteins. *Nucleic Acids Res* **42**: 13082–13095.

700 Paillard, S., and Strauss, F. (1991) Analysis of the mechanism of interaction of simian Ku protein with  
701 DNA. *Nucleic Acids Res* **19**: 5619–5624.

702 Paris, Ü., Mikkil, K., Tavita, K., Saumaa, S., Teras, R., and Kivisaar, M. (2015) NHEJ enzymes LigD and  
703 Ku participate in stationary-phase mutagenesis in *Pseudomonas putida*. *DNA Repair (Amst)* **31**: 11–  
704 18.

705 Pitcher, R.S., Brissett, N.C., and Doherty, A.J. (2007a) Nonhomologous end-joining in bacteria: a  
706 microbial perspective. *Annu Rev Microbiol* **61**: 259–282.

707 Pitcher, R.S., Green, A.J., Brzostek, A., Korycka-Machala, M., Dziadek, J., and Doherty, A.J. (2007b)  
708 NHEJ protects mycobacteria in stationary phase against the harmful effects of desiccation. *DNA*  
709 *Repair* **6**: 1271–1276.

- 710 Pitcher, R.S., Tonkin, L.M., Daley, J.M., Palmbo, P.L., Green, A.J., Velting, T.L., *et al.* (2006)  
711 Mycobacteriophage exploit NHEJ to facilitate genome circularization. *Mol Cell* **23**: 743–748.
- 712 Pitcher, R.S., Tonkin, L.M., Green, A.J., and Doherty, A.J. (2005) Domain structure of a NHEJ DNA  
713 repair ligase from *Mycobacterium tuberculosis*. *J Mol Biol* **351**: 531–544.
- 714 Płociński, P., Brissett, N.C., Bianchi, J., Brzostek, A., Korycka-Machata, M., Dziembowski, A., *et al.*  
715 (2017) DNA Ligase C and Prim-PolC participate in base excision repair in mycobacteria. *Nat Commun*  
716 **8** <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC5663958/>.
- 717 Popa, O., Hazkani-Covo, E., Landan, G., Martin, W., and Dagan, T. (2011) Directed networks reveal  
718 genomic barriers and DNA repair bypasses to lateral gene transfer among prokaryotes. *Genome*  
719 *Research* **21**: 599–609.
- 720 Raguse, M., Torres, R., Seco, E.M., Gándara, C., Ayora, S., Moeller, R., and Alonso, J.C. (2017) *Bacillus*  
721 *subtilis* DisA helps to circumvent replicative stress during spore revival. *DNA Repair (Amst)* **59**: 57–68.
- 722 Ramsden, D.A., and Asagoshi, K. (2012) DNA polymerases in nonhomologous end joining: are there  
723 any benefits to standing out from the crowd? *Environ Mol Mutagen* **53**: 741–751.
- 724 Ramsden, D.A., and Gellert, M. (1998) Ku protein stimulates DNA end joining by mammalian DNA  
725 ligases: a direct role for Ku in repair of DNA double-strand breaks. *EMBO J* **17**: 609–614.
- 726 Ricchetti, M., Fairhead, C., and Dujon, B. (1999) Mitochondrial DNA repairs double-strand breaks in  
727 yeast chromosomes. *Nature* **402**: 96–100.
- 728 Roberts, S.A., Strande, N., Burkhalter, M.D., Strom, C., Havener, J.M., Hasty, P., and Ramsden, D.A.  
729 (2010) Ku is a 5'-dRP/AP lyase that excises nucleotide damage near broken ends. *Nature* **464**: 1214–  
730 1217.
- 731 Setlow, B., and Setlow, P. (1996) Role of DNA repair in *Bacillus subtilis* spore resistance. *J Bacteriol*  
732 **178**: 3486–3495.
- 733 Shee, C., Gibson, J.L., Darrow, M.C., Gonzalez, C., and Rosenberg, S.M. (2011) Impact of a stress-  
734 inducible switch to mutagenic repair of DNA breaks on mutation in *Escherichia coli*. *Proc Natl Acad*  
735 *Sci USA* **108**: 13659–13664.
- 736 Shen, M., Zhang, H., Shen, W., Zou, Z., Lu, S., Li, G., *et al.* (2018) *Pseudomonas aeruginosa* MutL  
737 promotes large chromosomal deletions through non-homologous end joining to prevent  
738 bacteriophage predation. *Nucleic Acids Res* **46**: 4505–4514.
- 739 Shibata, A. (2017) Regulation of repair pathway choice at two-ended DNA double-strand breaks.  
740 *Mutation Research/Fundamental and Molecular Mechanisms of Mutagenesis* **803–805**: 51–55.
- 741 Shuman, S., and Glickman, M.S. (2007) Bacterial DNA repair by non-homologous end joining. *Nature*  
742 *Reviews Microbiology* **5**: 852–861.
- 743 Sinai, L., Rosenberg, A., Smith, Y., Segev, E., and Ben-Yehuda, S. (2015) The molecular timeline of a  
744 reviving bacterial spore. *Mol Cell* **57**: 695–707.
- 745 Sinha, K.M., Stephanou, N.C., Gao, F., Glickman, M.S., and Shuman, S. (2007) Mycobacterial UvrD1 is  
746 a Ku-dependent DNA helicase that plays a role in multiple DNA repair events, including double-strand  
747 break repair. *J Biol Chem* **282**: 15114–15125.

- 748 Sinha, K.M., Unciuleac, M.-C., Glickman, M.S., and Shuman, S. (2009) AdnAB: a new DSB-resecting  
749 motor–nuclease from mycobacteria. *Genes Dev* **23**: 1423–1437.
- 750 Stephanou, N.C., Gao, F., Bongiorno, P., Ehrt, S., Schnappinger, D., Shuman, S., and Glickman, M.S.  
751 (2007) Mycobacterial Nonhomologous End Joining Mediates Mutagenic Repair of Chromosomal  
752 Double-Strand DNA Breaks. *J Bacteriol* **189**: 5237–5246.
- 753 Vega, M. de (2013) The minimal *Bacillus subtilis* nonhomologous end joining repair machinery. *PLoS*  
754 *ONE* **8**: e64232.
- 755 Vlašić, I., Mertens, R., Seco, E.M., Carrasco, B., Ayora, S., Reitz, G., *et al.* (2014) *Bacillus subtilis* RecA  
756 and its accessory factors, RecF, RecO, RecR and RecX, are required for spore resistance to DNA  
757 double-strand break. *Nucleic Acids Res* **42**: 2295–2307.
- 758 Vries, E. de, Driel, W. van, Bergsma, W.G., Arnberg, A.C., and Vliet, P.C. van der (1989) HeLa nuclear  
759 protein recognizing DNA termini and translocating on DNA forming a regular DNA-multimeric protein  
760 complex. *J Mol Biol* **208**: 65–78.
- 761 Walker, J.R., Corpina, R.A., and Goldberg, J. (2001) Structure of the Ku heterodimer bound to DNA  
762 and its implications for double-strand break repair. *Nature* **412**: 607–614.
- 763 Wang, S.T., Setlow, B., Conlon, E.M., Lyon, J.L., Imamura, D., Sato, T., *et al.* (2006) The Forespore Line  
764 of Gene Expression in *Bacillus subtilis*. *Journal of Molecular Biology* **358**: 16–37.
- 765 Ward, J.F. (1994) The Complexity of DNA Damage: Relevance to Biological Consequences.  
766 *International Journal of Radiation Biology* **66**: 427–432.
- 767 Weller, G.R., Kysela, B., Roy, R., Tonkin, L.M., Scanlan, E., Della, M., *et al.* (2002) Identification of a  
768 DNA nonhomologous end-joining complex in bacteria. *Science* **297**: 1686–1689.
- 769 Zhang, X., Chen, W., Zhang, Y., Jiang, L., Chen, Z., Wen, Y., and Li, J. (2012) Deletion of *ku* homologs  
770 increases gene targeting frequency in *Streptomyces avermitilis*. *J Ind Microbiol Biotechnol* **39**: 917–  
771 925.
- 772 Zhou, Y., Chen, T., Zhou, L., Fleming, J., Deng, J., Wang, X., *et al.* (2015) Discovery and  
773 characterization of Ku acetylation in *Mycobacterium smegmatis*. *FEMS Microbiol Lett* **362**  
774 <https://academic-oup-com.bases-doc.univ-lorraine.fr/femsle/article/362/6/fnu051/580930>.  
775 Accessed June 27, 2018.
- 776 Zhu, H., and Shuman, S. (2005a) A primer-dependent polymerase function of *Pseudomonas*  
777 *aeruginosa* ATP-dependent DNA ligase (LigD). *J Biol Chem* **280**: 418–427.
- 778 Zhu, H., and Shuman, S. (2005b) Novel 3′-ribonuclease and 3′-phosphatase activities of the bacterial  
779 non-homologous end-joining protein, DNA ligase D. *J Biol Chem* **280**: 25973–25981.
- 780 Zhu, H., and Shuman, S. (2007) Characterization of *Agrobacterium tumefaciens* DNA ligases C and D.  
781 *Nucleic Acids Res* **35**: 3631–3645.
- 782 Zhu, H., and Shuman, S. (2010) Gap Filling Activities of *Pseudomonas* DNA Ligase D (LigD) Polymerase  
783 and Functional Interactions of LigD with the DNA End-binding Ku Protein. *J Biol Chem* **285**: 4815–  
784 4825.



785 Zhu, H., Wang, L.K., and Shuman, S. (2005) Essential constituents of the 3'-phosphoesterase domain  
786 of bacterial DNA ligase D, a nonhomologous end-joining enzyme. *J Biol Chem* **280**: 33707–33715.

787

788

Submitted Version

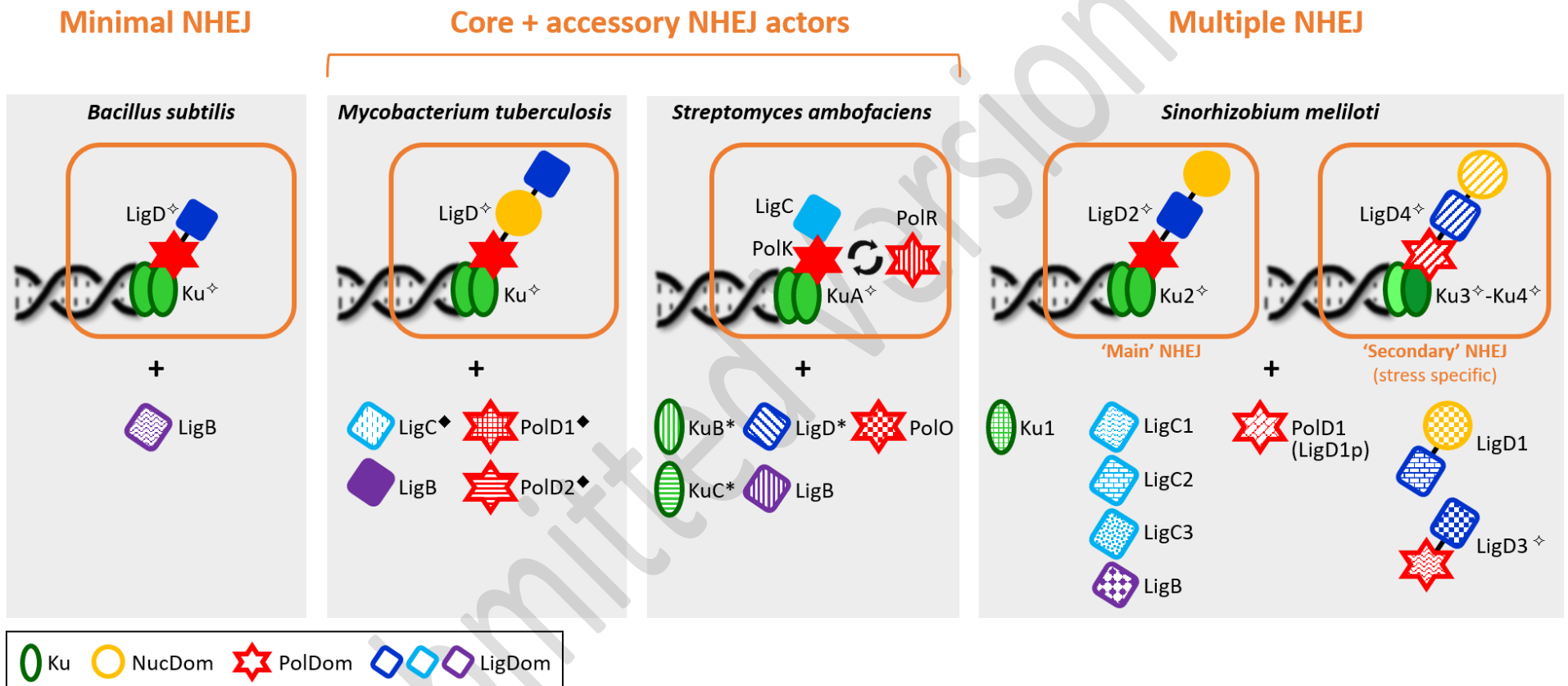


Figure 1: Repertoires of NHEJ presumed actors in model bacteria from minimal NHEJ to the more complex situations. Ku, LigDom (ATP-dependent), NucDom, PolDom identified in the genome of the bacterial species (as described in the literature) are highlighted whatever present within a single polypeptide or as standalone domains. Proteins having the same name in different organisms are not necessarily encoded by orthologous genes. <sup>◇</sup> indicate actors which have been experimentally involved in DSB repair. <sup>◆</sup> in *M. tuberculosis* denote accessory actors involved in NHEJ, based on studies of their homologs in *M. smegmatis* coding for different LigD with mutated domain (see text). In *S. ambofaciens*, accessory actors involved in response to DNA damage are highlighted by an asterisk (Hoff *et al.*, 2016, 2018). For the sake of clarity, only one dimer of Ku is shown at DNA ends and the ability of Ku to thread inward the DNA molecule has not been represented.