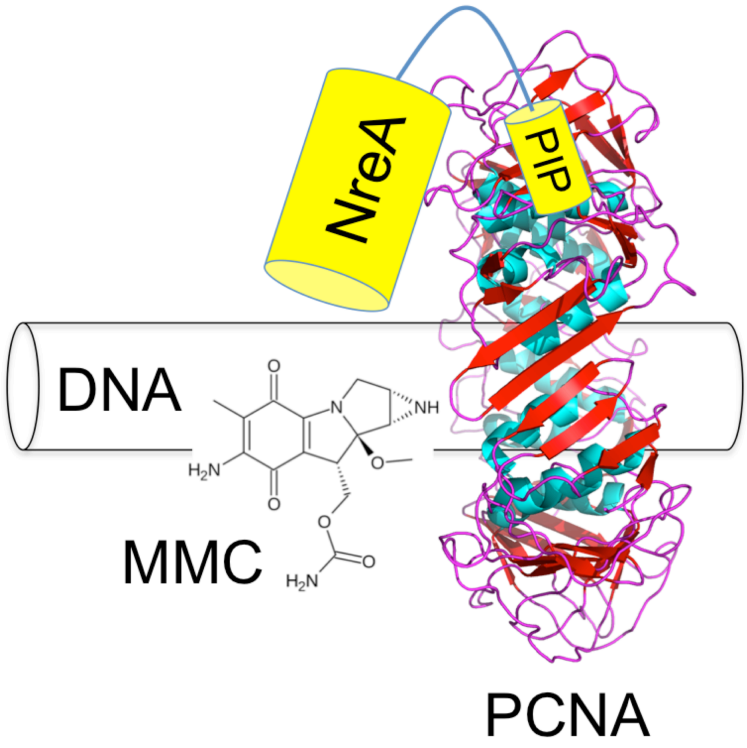
A novel archaeal DNA repair factor that acts with the UvrABC system to repair mitomycin C-induced DNA damage in a PCNA-dependent manner

Xavier Giroux and Stuart A. MacNeill

Abbreviated Summary

Molecular genetic analysis in the halophilic euryarchaeal organism *Haloferax volcanii* is used to demonstrate a role for NreA, a member of a previously uncharacterised but highly conserved family of archaeal proteins, in the cellular response to DNA damage caused by the DNA crosslinker mitomycin C. NreA function appears dependent on its ability to interact with the sliding clamp PCNA via a C-terminal PIP (PCNA interacting protein) motif.



A novel archaeal DNA repair factor that acts with the UvrABC system to repair mitomycin C-induced DNA damage in a PCNA-dependent manner

Xavier Giroux and Stuart A. MacNeill *

Biomedical Sciences Research Complex
School of Biology
University of St Andrews
North Haugh
St Andrews
Fife KY16 9ST
UK.

* For correspondence:
E-mail: stuart.macneill@st-andrews.ac.uk
Telephone: +44 01334 46 7268

Running title: Novel archaeal PCNA-dependent DNA repair factor

Keywords: DNA repair, sliding clamp, double-strand breaks, archaea, PCNA

Summary

The sliding clamp PCNA plays a vital role in a number of DNA repair pathways in eukaryotes and archaea by acting as a stable platform onto which other essential protein factors assemble. Many of these proteins interact with PCNA via a short peptide sequence known as a PIP (PCNA interacting protein) motif. Here we describe the identification and functional analysis of a novel PCNA interacting protein NreA that is conserved in the archaea and which has a PIP motif at its C-terminus. Using the genetically tractable euryarchaeon *Haloferax volcanii* as a model system, we show that the NreA protein is not required for cell viability but that loss of NreA (or replacement of the wild-type protein with a truncated version lacking the C-terminal PIP motif) results in an increased sensitivity to the DNA damaging agent mitomycin C (MMC) that correlates with delayed repair of MMC-induced chromosomal DNA damage monitored by pulsed-field gel electrophoresis (PFGE). Genetic epistasis analysis in *Hfx. volcanii* suggests that NreA works together with the UvrABC proteins in repairing DNA damage resulting from exposure to MMC. The wide distribution of NreA family members implies an important role for the protein in DNA damage repair in all archaeal lineages.

Introduction

DNA damage presents a severe and ongoing threat to the overall stability of the genome and the integrity of the genetic information. In all three domains of life DNA damage is countered by an extensive overlapping network of damage surveillance, repair and mitigation processes. Central to these processes are a large number of proteins that have evolved to recognize various types of DNA lesion and to catalyse their repair. In humans, mutations in several DNA repair pathways have been shown to result in predisposition to cancer. For example, inactivation of nucleotide excision repair (NER) enzymes (Scharer, 2013) is the most common cause of xeroderma pigmentosum (XP), an autosomal recessive disorder associated with an extremely high incidence of skin cancer. NER is the major pathway for the repair of UV damage in cells and XP patients are acutely sensitive to UV light due to their inability to repair UV-induced lesions (Scharer, 2013). Understanding how DNA is repaired is of great importance for human health and as many of the basic mechanisms of DNA repair are conserved across all three domains of life, simple microbial systems can be used as models in the quest for a detailed understanding of these mechanisms.

The sliding clamp PCNA (proliferating cell nuclear antigen) plays a central role in many aspects of DNA metabolism in eukaryotic and archaeal cells (Maga & Hubscher, 2003, Moldovan *et al.*, 2007). PCNA is a trimeric ring-shaped oligomer that is loaded onto double-stranded DNA in an ATP-dependent reaction catalysed by the clamp loader complex, replication factor C (Yao & O'Donnell, 2012). Once loaded onto DNA, PCNA serves as a stable platform onto which other proteins can assemble (Tsurimoto, 2006). While PCNA itself has no enzymatic activity, the majority of its binding partners are enzymes with DNA-directed activities, such as DNA polymerases, nucleases, methylases, glycosylases and ligases. PCNA partners participate in a wide range of processes and pathways, including chromosomal DNA replication, various DNA repair pathways (such as nucleotide excision repair, base excision repair, mismatch repair, etc.), recombination and cell cycle control (Tsurimoto, 2006).

In eukaryotes and most archaeal lineages, PCNA is homotrimeric, with heterotrimeric forms being seen in a sub-set of archaea only (Dieckman *et al.*, 2012). In both homotrimeric or heterotrimeric PCNA complexes, each monomer is composed of two structurally similar domains and the three monomers are assembled in a head-to-tail fashion, giving the clamp structure overall six-fold symmetry (Krishna *et al.*, 1994, Gulbis *et al.*, 1996). Bacterial sliding clamps also have a ring-shaped architecture but are dimeric, with each monomer being made up of three related domains to maintain

the six-fold symmetry (Kong *et al.*, 1992). The internal surface of PCNA is mainly formed by α-helices and is positively charged. This promotes interactions with DNA but still allows the PCNA trimer to slide along DNA. In contrast, the outer surface of PCNA is composed of β-sheets and is negatively charged. This region is involved in protein-protein interactions (Krishna *et al.*, 1994, Gulbis *et al.*, 1996).

Many of the proteins that interact with PCNA do so via a conserved short peptide sequence called a PCNA interacting peptide (PIP) motif that binds tightly to a hydrophobic pocket on the surface of the PCNA trimer (Warbrick, 1998). The PIP motif has consensus sequence Q x x (*a*) x x (*h h*), where *a* represents amino acids with small hydrophobic side chains (usually I or L) and *h* represents amino acids with aromatic hydrophobic side chains (F, Y, W), and is often located at the extreme N- or C-terminus of the interacting protein (Warbrick, 1998). Structures of eukaryotic and archaeal PCNA clamps in complex with PIP motif peptides from a variety of proteins have been determined by X-ray crystallography (Gulbis *et al.*, 1996), allowing PCNA-PIP interactions to be visualised in great detail (Winter & Bunting, 2012).

In this report we describe functional analysis of a novel conserved archaeal PCNA interacting protein using the halophilic euryarchaeon *Haloferax volcanii* as a model system. The haloarchaea offer an attractive model for studies of archaeal DNA replication, repair and recombination owing to the ease with which representative species can be manipulated genetically (Farkas *et al.*, 2013). *Hfx. volcanii* in particular has proved a particularly successful model, with the functions of a number of key replication, repair and recombination factors already characterized genetically, including single-stranded DNA binding proteins (Skowyra & MacNeill, 2012, Stroud *et al.*, 2012), the MCM helicase (Kristensen *et al.*, 2014), the Rad50-Mre11 complex (Delmas *et al.*, 2013, Delmas *et al.*, 2009), the Holliday junction resolvase Hjc (Lestini *et al.*, 2010, Lestini *et al.*, 2013), ATP and NAD-dependent DNA ligases (Poidevin & MacNeill, 2006, Zhao *et al.*, 2006), and the Hef, RNaseHII and Fen1 nucleases (Lestini *et al.*, 2010, Meslet-Cladiere *et al.*, 2007). On the basis of their sequence similarity to biochemically well-characterised homologues in other archaeal species (see Meslet-Cladiere *et al.*, 2007, for example), the latter two *Hfx. volcanii* proteins, RNaseHII and Fen1, appear likely to bind to PCNA via conserved PIP motifs, although this has not been confirmed biochemically. On the same basis (orthology with known PCNA binding PIP motif-containing proteins in other archaeal species), six other *Hfx. volcanii* proteins (excluding the NreA protein described in this report) can be postulated to bind to PCNA via PIP motifs, including three DNA polymerases and a DNA helicase (discussed further below). In each case the putative PIP motifs are located at the extreme C-termini of the proteins. While the three-dimensional

structure of *Hfx. volcanii* PCNA has also been solved and the putative PIP motif binding pocket identified (Morgunova *et al.*, 2009, Winter *et al.*, 2009), no biochemical analysis of haloarchaeal PCNA-PIP interactions has been reported, nor has the *in vivo* importance of PCNA binding via PIP motifs been probed genetically.
The highly conserved nature of the PIP motif allows it to be used as a tool to identify novel PCNA binding proteins in bioinformatic screens. Here, we report the identification and characterization of a previously unstudied archaeal protein (NreA) carrying a C-terminal PIP motif sequence. Nre proteins are found encoded by almost all archaeal lineages in one or two copies per genome and the C-terminal PIP motif is a conserved feature of these proteins. Using the *Hfx. volcanii* as a model, we show that NreA is non-essential for cell viability but that cells lacking NreA are sensitive to DNA damage induced by treatment with mitomycin C (MMC). This sensitivity is associated with a significant delay in the repair of damaged chromosomes (visualised by PFGE) following MMC treatment. Genetic epistasis analysis shows that NreA acts in a different pathway to the repair factors Rad50-Mre11, Hjc, Hef and UvrD in MMC repair and points instead to a functional relationship with homologues of the bacterial UvrABC proteins implicated in nucleotide excision repair. Finally, we show that the phenotypes resulting from complete loss of NreA function are mirrored by deletion of the C-terminal PIP motif also, implying that NreA function is dependent on its ability to bind to PCNA. Taken together, our results identify a highly conserved archaeal PCNA interacting protein as an important new player in the repair of DNA damage in the archaea.

Results

Identification of a new family of PCNA interacting proteins

The presence of one or more PIP motifs in many PCNA interacting proteins allows this sequence to be used to identify novel factors potentially involved in DNA metabolism (replication, repair and recombination). Using the PIP motifs of likely *Hfx. volcanii* PCNA binding proteins (i.e. proteins with known PIP motif-containing PCNA binding orthologues in other archaeal species, as discussed in the Introduction) as the query sequences in BLAST searches identified several potential novel PIP motif-containing proteins encoded by this organism, including a member of a previously uncharacterized family of archaeal proteins that carry a PIP motif at their C-terminal end (Table 1). These proteins (designated the Nre family) are unique to the archaea: no bacterial, eukaryotic or viral Nre homologues are apparent. In a previous study, an

Archaeoglobus fulgidus Nre protein was identified using the yeast two-hybrid system as interacting with PCNA, although no attempt was made to further characterise its cellular function (Motz *et al.*, 2002).

The Nre proteins, which are ~ 400 amino acids long, belong to COG1602 and arCOG04269 in the COG (Clusters of Orthologous Genes) and archaeal COG (arCOG) databases, respectively (Tatusov *et al.*, 2003, Tatusov *et al.*, 2001, Tatusov *et al.*, 2000, Wolf *et al.*, 2012, Makarova *et al.*, 2007). Note that in the Pfam database (Finn *et al.*, 2014), N- and C-terminal regions of the Nre proteins are designated as two separate domains, DUF650 (domain of unknown function 650) and DUF651, due to an error in the published *A. fulgidus* genome sequence (see Supplementary information for further details). Fig. 1A presents a schematic overview of the domain organisation of the Nre proteins from selected archaeal species (see also Table 1). Genes encoding Nre proteins are found in almost all archaeal species whose genomes have been sequenced. Exceptions include *Methanocaldococcus jannaschii* (although the protein is present in other *Methanocaldococcus* species) and the obligate symbiont *Nanoarchaeum equitans*.

While most archaeal organisms encode only a single Nre protein, some encode two, from separate genes, in which case we have designated these NreA and NreB (Fig. 1A, see below). The PIP motif is highly conserved in the NreA proteins (Fig. 1B) but is typically less well conserved (and often undetectable) in the NreB proteins.

In addition to possessing a C-terminal PIP motif, many Nre proteins have a predicted C4 metal binding domain at their N-terminal end (Fig. 1C). The C4 metal binding domain could either be a zinc finger DNA binding domain or an [4Fe-4S] iron-sulphur cluster. In the past few years, iron-sulphur clusters have been found in several eukaryotic and archaeal DNA replication and repair enzymes, such as the DNA repair helicases XPD and FANCJ (Rudolf *et al.*, 2006) as well as the bacterial spore photoproduct lyase (SplB, see below) (Rebeil *et al.*, 1998). The functions of these iron-sulphur clusters are not clear, but it has been proposed that they may act during the search of DNA damage by acting as a redox cofactor for DNA-mediated charge transport (Boal *et al.*, 2007). The C4 metal binding domain cannot be essential for Nre function, however, as it is absent in some species' Nre proteins (Fig. 1A).

To gain further insight into potential roles for the Nre proteins, we examined the genomic context of Nre-encoding genes in diverse species (Fig. 1D, Table 1). In both euryarchaeal and crenarchaeal organisms, genes encoding NreB proteins were found adjacent to a gene encoding a protein with significant similarities with the bacterial SplB enzyme (Chandor *et al.*, 2006, Buis *et al.*, 2006, Rebeil & Nicholson, 2001, Slieman *et al.*, 2000, Rebeil *et al.*, 1998). SplB (SP lyase) is responsible for the

repair of a unique type of DNA damage (the intrastrand thymine dimer 5-thyminyl-5,6-dihydrothymine, known as spore photoproduct or SP) found in UV-irradiated bacterial spore DNA. The evolutionarily conserved linkage between the Nre and SplB-like genes suggests that the corresponding proteins could be functionally linked. It should be noted, however, that there are no reports of SP formation in archaeal organisms to date, and thus the substrate of the archaeal SplB-like enzymes is unknown.

Hfx. volcanii NreA is not essential but is needed for efficient repair of mitomycin C-induced DNA damage

To investigate the function of Nre proteins in archaea, we exploited the reverse genetic tools available for *Hfx. volcanii.* This organism encodes a single Nre family protein from the *HVO_0734* gene, hereafter referred to as *nreA* (Fig. 2A). The *nreA* ORF is located 85 bp downstream from the 3' end of a gene (*tfb4*) encoding a transcription initiation factor B (TFB) protein. A putative TATA box (5'-TTCATAT-3') is located approximately 25-30 bp upstream of the *nreA* start codon (Palmer & Daniels, 1995), indicating that the *nreA* mRNA is likely to be leaderless, in common with many other *Hfx. volcanii* mRNA transcripts (Hering *et al.*, 2009). 36 bp downstream of *nreA* lies *HVO_0735* which encodes a short (57 residues) protein of unknown function that is conserved in a few other haloarchaeal organisms only. Expression of *HVO_0735* is likely driven by a TATA box located just beyond the *nreA* stop codon. The *nreA* mRNA is therefore likely to be monocistronic.

The NreA protein is predicted to be 420 amino acids in length and like other haloarchaeal Nre proteins, NreA lacks the N-terminal C4 metal binding motif (Figs. 1A, 2A). The PIP motif has the sequence QTDIFDFA (conserved residues underlined) and is located at the immediate C-terminus of the protein, a characteristic of all known PIP motifs in *Hfx. volcanii* (Fig. 2B).

To delete the *nreA* gene from the chromosome of *Hfx. volcanii* cells, we used the pop-in/pop-out method, initially replacing *nreA* with the selectable marker *trpA* (see Experimental procedures for details). Cells deleted for *nreA* were readily obtained by this method, indicating that the NreA protein is not required for cell viability in *Hfx. volcanii*. Indeed, when cultured under standard conditions, growth of Δ*nreA* is indistinguishable from wild-type (Fig. 2C).

Next we examined whether Δ*nreA* cells displayed increased sensitivity to DNA damaging agents, as would be anticipated if NreA has role in DNA repair in *Hfx. volcanii*. Wild-type and Δ*nreA* cells were exposed to three diverse DNA damaging

agents: ultraviolet (UV) light, methyl methanesulphonate (MMS) and mitomycin C (MMC). UV exposure causes formation of two types of DNA lesion, cyclobutane pyrimidine dimers (CPDs) and pyrimidine 6-4 photoproducts (6-4PPs), whereas MMS methylates DNA, predominantly converting guanine to 7-methylguanine and adenine to 3-methyladenine (Beranek, 1990). MMC reacts with guanine to produce three types of MMC damage: MMC-mono-dG-adducts, intrastrand dG-MMC-dG biadducts and interstrand dG-MMC-dG crosslinks (Bargonetti *et al.*, 2010, Tomasz, 1995). No increase in sensitivity, relative to wild-type cells, was seen with UV or MMS (data not shown, see Experimental procedures for doses tested). However, we found that Δ*nreA* cells displayed a moderate increase in sensitivity to MMC (Fig. 2D) and explored this further.

NreA is needed for efficient repair of double strand breaks

The above results indicate a role for NreA in the repair of MMC-induced DNA damage. To investigate the sensitivity of Δ*nreA* cells further, we monitored chromosome damage and repair in *Hfx. volcanii* using a pulsed field gel electrophoresis (PFGE) assay. Mid-exponential cultures of wild-type *Hfx. volcanii* and Δ*nreA* strains growing in YPC medium were exposed to MMC at final concentration of 0.5 μg/ml for 1 hour. The MMC was then removed and the cells were allowed to recover for 10 hours. Samples for PFGE were taken prior to MMC exposure, immediately afterwards, and every two hours after MMC was removed. Samples were prepared for PFGE as described in the Experimental procedures.
As can be seen in Fig. 3, the chromosomes of *Hfx. volcanii* are broken by exposure to MMC. Double-strand breaks (DSBs) in DNA are a known consequence of the inter-strand crosslink repair in replicating cells. This is visible in PFGE as the disappearance of the DNA fragments, and the appearance of a DNA smear at the bottom of the gel. In wild-type *Hfx. volcanii*, repair of the DSBs is first detectable after 4 hours, and is completed within 6 to 8 hours. In contrast, in Δ*nreA* cells, little or no repair is seen 10 hours after MMC treatment. The NreA protein is therefore required for efficient repair of DNA double-strand breaks in *Hfx. volcanii.*

NreA acts together with the UvrABC system to repair MMC-induced DNA damage

In order to investigate whether NreA performs its function in association with any of the previously analysed DNA repair factors in *Hfx. volcanii*, we performed genetic epistasis analysis. To do this, we again deleted the *nreA* gene but this time in cells

lacking proteins involved in various aspects of DNA repair: Mre11 and Rad50, the helicase-nuclease Hef, the Holliday junction resolvase Hjc, and homologues of the bacterial NER proteins UvrA - UvrD. In *Hfx. volcanii*, Mre11 and Rad50 play a role in DNA double-strand break repair (DSBR), apparently by restraining homologous recombination to allow other repair pathways to act, Hef and Hjc have been implicated in replication fork restart, and UvrA, UvrB and UvrC are required for efficient repair of UV damaged DNA (Delmas *et al.*, 2009, Lestini *et al.*, 2010). No function has yet been assigned to the putative helicase UvrD.

As shown in Figs. 4A-4F, *Hfx. volcanii* Δ*mre11* Δ*rad50*, Δ*hef*, Δ*hjc*, Δ*uvrA*, Δ*uvrB* and Δ*uvrC* strains are all more sensitive to MMC than the wild-type, confirming previous results with Δ*hef* and Δ*hjc* (Lestini *et al.*, 2010) and demonstrating for the first time the involvement of the Mre11-Rad50 and UvrABC systems in MMC-induced damage repair in the haloarchaea. No increase in sensitivity was seen in cells lacking UvrD (Fig. 4G), indicating that this protein is non-essential for MMC-induced damage repair, just as it is non-essential for UV damage repair (Lestini *et al.*, 2010).

Deletion of *nreA* in Δ*mre11* Δ*rad50*, Δ*hef*, Δ*hjc* or Δ*uvrD* backgrounds leads to a further increase in MMC sensitivity (Figs. 4A-4C), implying that NreA's role in MMC-damage repair is distinct from those of Mre11-Rad50, Hef, Hjc or UvrD. However, in sharp contrast to these results, deletion of *nreA* in any of the *uvrA*, *uvrB* or *uvrC* backgrounds does not lead to an increase MMC sensitivity beyond that of the *uvrA*, *uvrB* and *uvrC* single mutants (Figs. 4D-4F), indicating that NreA acts together with UvrA, UvrB and UvrC to repair MMC-induced DNA damage in *Hfx. volcanii.*

In an attempt to further characterise the relationship between the NreA, UvrA, UvrB and UvrC proteins, we assessed the UV sensitivity of wild-type, Δ*nreA* and Δ*nreA uvrA* strains. As shown in Fig. 4H, the Δ*nreA* strain is no more sensitive to UV irradiation than wild-type, and the double mutant Δ*nreA* Δ*uvrA* is no more sensitive than the single Δ*uvrA* mutant. Thus, NreA is not part of the NER system that repairs UV induced DNA damage.

NreA function is dependent on the presence of the PIP motif

The results above indicate that the NreA protein functions with UvrABC in the repair of MMC-induced chromosomal DNA damage in *Hfx. volcanii*. The presence of a PIP motif at the C-terminal end of NreA could reflect a requirement for PCNA binding during the repair reaction, for localization of NreA to sites of DNA damage for example. To investigate the role of PCNA binding, we constructed strains expressing a truncated NreA protein lacking the C-terminal 10 amino acids encompassing the

entire PIP motif (see Fig. 1C and Experimental procedures for details), and exposed these *nreA-CΔ10* cells to MMC. As shown in Fig. 2C, the viability of *nreA-CΔ10* cells plated on increasing concentrations of MMC is essentially indistinguishable to that of Δ*nreA* cells. We also deleted the NreA PIP motif in the following backgrounds: Δ*mre11* Δ*rad50*, Δ*hef*, Δ*hjc*, Δ*uvrA*, Δ*uvrB*, Δ*uvrC* and Δ*uvrD*. In each case, the Δ*nreA-CΔ10* strain phenocopied the corresponding Δ*nreA* strain (Fig. 4A - 4G). Finally, we investigated the effects of deletion of the NreA PIP motif on recovery from MMC treatment by PFGE (Figure 3). As with Δ*nreA* cells, *nreA-CΔ10* cells displayed increased chromosome fragmentation and delayed recovery relative to wild-type *Hfx. volcanii*. We conclude that the presence of the putative PIP motif, and most likely therefore interaction with PCNA, is essential for NreA function in the repair process.

Discussion

Sliding clamps play a central role in many aspects of DNA metabolism in all three domains of life. In eukaryotes, the sliding clamp PCNA is a key component of the chromosome replication machinery and is additionally involved in DNA damage repair, cell cycle control, sister chromatid cohesion and chromatin remodeling (De Biasio & Blanco, 2013, Hedglin *et al.*, 2013). These roles are mediated largely through direct protein-protein interactions with a wide range of protein factors including a number of enzymes that act directly on DNA substrates, such as polymerases, ligases, nucleases, methyltransferases and glycosylases (Tsurimoto, 2006). Archaea also possess a PCNA clamp and this too has been implicated in diverse DNA transactions through protein-protein interactions. Here, we report the identification and functional characterisation of a new family of PCNA interacting proteins (designed Nre) involved in DNA repair in the archaea (Fig. 1). Using the genetically tractable haloarchaeal organism *Hfx. volcanii* as a model, we have shown that the NreA protein is non-essential for cell viability (NreA is the sole Nre family protein encoded by *Hfx. volcanii*) but that Δ*nreA* cells are hypersensitive to the DNA damaging agent mitomycin C (MMC) and are slow to repair chromosomal DNA following MMC treatment (Fig. 2, Fig. 3). Genetic epistasis analysis indicates that NreA acts together with the UvrABC proteins in the repair of MMC damage (Fig. 4).

The previously uncharacterised Nre proteins were initially identified by BLAST searching (Altschul & Koonin, 1998, Altschul *et al.*, 1997) using the conserved PIP (PCNA interacting protein) motif (Warbrick, 1998) as the query sequence. Nre proteins are encoded by one or two genes in all almost archaea whose genomes have been sequenced. In many cases genes encoding Nre proteins are found

adjacent to (or overlapping with) those encoding archaeal homologues of spore photoproduct (SP) lyase SplB (Chandor *et al.*, 2006, Buis *et al.*, 2006, Rebeil & Nicholson, 2001, Slieman *et al.*, 2000, Rebeil *et al.*, 1998) (Fig. 1D, Table 1). SplB is responsible for the repair of the unusual intrastrand thymine dimer 5-thyminyl-5,6-dihydrothymine found in UV-irradiated bacterial spore DNA and could play a similar role in archaea, although it should be noted that there are no reports of formation of SP-like lesions in archaea and that the substrates of the archaeal SplB proteins remain unknown. Nevertheless, we initially hypothesized that Nre proteins, possibly associated with the SplB homologues when present, could be involved in the repair UV-induced bulky adducts, either classical cyclobutane pyrimidine dimers (CPDs) and/or pyrimidine (6-4) photoproducts (PPs) or more complex types of damage such as spore photoproducts. *Hfx. volcanii* Δ*nreA* cells do not display increased sensitivity to UV damage compared to wild-type (Fig. 4H), however, so a major role in repair UV-induced damage is unlikely.

Next we tested the sensitivity of Δ*nreA* strains to mitomycin C (MMC). MMC reacts with guanine to form three types of MMC damage lesion: MMC-mono-dG-adducts, intrastrand dG-MMC-dG biadducts and interstrand dG-MMC-dG crosslinks (Bargonetti *et al.*, 2010, Tomasz, 1995). Δ*nreA* cells display an moderate increase in sensitivity to MMC (Fig. 2D), indicating a role for NreA function in the repair of MMC-induced DNA lesions. A similar increase in sensitivity was seen in cells expressing NreA lacking the C-terminal PCNA binding PIP motif, implying that the function of NreA in repairing MMC damage is dependent on its interaction with PCNA (Fig. 2D). The increased sensitivity of Δ*nreA* and *nreA-C*Δ*10* strains to MMC correlates with delayed ability to repair chromosomal DNA damage as judged by PFGE (Fig. 3).

These observations raise the question of which type of MMC-induced damage lesion is repaired by NreA? In bacterial and eukaryotic systems, MMC mono-adducts and intrastrand crosslinks are removed by the NER pathway, while repair of interstrand crosslinks (ICLs) requires more complex mechanisms and often involves multiple different pathways working together. Little is known of how this occurs in archaea, however. The only archaeal proteins that have been shown to be involved in this repair pathway are Hef and Hjc but their exact functions remain unknown (Lestini *et al.*, 2010, Lestini *et al.*, 2013). Other bacterial or eukaryotic proteins known to be involved in ICL repair are present in archaea, including the Mre11-Rad50 complex (Delmas *et al.*, 2013, Delmas *et al.*, 2009) and the NER system, which in the haloarchaea (Crowley *et al.*, 2006, Lestini *et al.*, 2010) involves homologues of the bacterial UvrA, UvrB and UvrC proteins, the genes for which where presumably acquired by lateral gene transfer from bacteria. Our results confirm that all these

factors are involved in the repair of MMC-induced DNA damage in *Hfx.volcanii* (Fig. 4).

In eukaryotic cells, the MRN (Mre11-Rad50-Nbs1) complex is known to be involved in DNA double-strand break (DSB) repair through homologous recombination (HR). In contrast, *Hfx. volcanii* Mre11-Rad50 restrains HR, instead of promoting it through DNA end resection (Delmas *et al.*, 2009). Moreover, absence of the Mre11-Rad50 complex increases resistance to several types of DNA damage (UV, phleomycin, γ-irradiation and MMS). The higher sensitivity of the Δ*rad50* Δ*mre11* double deletion to MMC (Fig. 4A) points to a different role for the Mre11–Rad50 complex in the repair of MMC-induced damage. It has been shown that the MRX (Mre11-Rad50-Xrs2) complex is needed for maintaining the stability of stalled replication forks in eukaryotes (Tittel-Elmer *et al.*, 2009). If MMC leads to replication fork stalling in *Hfx. volcanii* and Mre11–Rad50 function is required to prevent fork collapse under these conditions, then this could explain the increased sensitivity of Δ*rad50* Δ*mre11* cells to MMC observed here.

Regarding the UvrABC system, this is known to be involved in the repair of MMC-induced DNA damage in other systems (Kisker *et al.*, 2013). NER can repair mono-adducts and intrastrand crosslinks, which are similar to CPD or PPs, and also ICLs. The results presented here confirm that UvrABC proteins are needed for efficient repair of MMC-induced DNA damage in *Hfx. volcanii* (Fig. 4D-F). Cells lacking any one of the three *uvrA*, *uvrB* or *uvrC* genes are significantly more sensitive to MMC than either wild-type or Δ*nreA* cells (Fig. 4D-F). Strikingly, the double mutants Δ*nreA* Δ*uvrA*, Δ*nreA* Δ*uvrB* and Δ*nreA* Δ*uvrC* are no more sensitive than the Δ*uvrA*, Δ*uvrB* and Δ*uvrC* single mutants, suggesting that NreA is involved in the same MMC damage repair pathway as UvrABC, such that in the absence of UvrABC function the presence or absence of NreA is irrelevant. Note, however, that Δ*nreA* cells are not UV sensitive, unlike Δ*uvrA*, Δ*uvrB* and Δ*uvrC* single mutants, and that a Δ*nreA* Δ*uvrA* double mutant is no more sensitive than the Δ*uvrA* single mutant (Fig. 4H). These results indicate that NreA does not play a general role in UvrABC-mediated NER but could suggest that NreA functions in a step specific to particular types of DNA damage.

In addition to UvrABC, *Hfx. volcanii* also encodes a homologue of the bacterial UvrD helicase. In bacteria, UvrD plays important roles in both NER and also in mismatch repair (MMR), the main pathway correcting DNA sequence errors arising from nucleotide insertion, deletion and misincorporation during replication and recombination (Kuper & Kisker, 2013). Recent results indicate that in a specific sub-pathway of NER called transcription-coupled DNA repair (TCR), UvrD acts by binding

to RNA polymerase (RNAP) and inducing RNAP backtracking, allowing NER enzymes to gain access to DNA lesions that would otherwise be shielded by the stalled polymerase (Brosh, 2014, Epshtein, 2015, Epshtein *et al.*, 2014). In MMR, UvrD acts to unwind DNA in the vicinity of the mismatch, thereby exposing single-stranded DNA to nuclease cleavage and subsequent removal (Kuper & Kisker, 2013). Whether one, both or neither of these roles is conserved in *Hfx. volcanii* cells remains to be seen.

We show here that, in contrast to what is observed with Δ*uvrA*, Δ*uvrB* and Δ*uvrC* single mutants, *Hfx. volcanii* Δ*uvrD* cells are no more sensitive to MMC than wild-type (Fig. 4G). It has previously been shown that Δ*uvrD* cells are not especially sensitive to UV either (Lestini *et al.*, 2010). UvrD therefore either plays no role in the repair of UV and MMC-induced DNA damage in *Hfx. volcanii*, or it can be replaced by another protein with redundant activity. Interestingly, when both UvrD and NreA are absent, cells become hypersensitive to MMC (Fig. 4G), suggesting that that UvrD and NreA might share a redundant function in the repair of MMC-induced damage by the UvrABC system. It is tempting to speculate that NreA could contribute either to backtracking of archaeal RNAP stalled at MMC-induced lesions or in preventing the RNAP from becoming stalled at MMC-induced lesions in the first place, or that NreA could act downstream of the MutS and MutL proteins (*Hfx. volcanii* encodes two of each) in MMR-type pathway that can act to remove MMC lesions.

As shown previously (Lestini *et al.*, 2010), strains lacking Hef and Hjc are more sensitive to MMC than wild-type (Fig. 4B, 4C). These authors proposed that Hef and Hjc act at replication fork stalled by MMC-induced ICLs, and that Hef could either reverse the replication using its helicase activity, allowing the Hjc nuclease to resolve the resulting Holliday junction, or that Hef could directly cut the fork leading to a double strand break. NreA does not appear to be part of this mechanism, as Δ*nreA* Δ*hjc* and Δ*nreA* Δ*hef* cells are more sensitive than the Δ*hjc* and Δ*hef* single mutants. However, there is a link between all these proteins. As shown in the PFGE assays, NreA is involved in the repair of the DSBs caused by MMC treatment (Fig. 3). These double strand breaks are likely to occur mainly at replication forks as the replication machinery encounters an ICL. At the point, in the absence of Hef, fork reversal may not be possible and thus the propensity for DSB formation will increase. If the increase in DSB formation occurs in cells lacking NreA, where DSB repair is reduced, then the overall effect will be a synergistic increase in MMC sensitivity, as observed (Fig. 4C).

With regard to Hjc, this protein is believed to be responsible for resolution of Holliday junctions formed during DNA double strand break repair by homologous

recombination (HR). The results shown in Fig. 4B indicate a role for Hjc in repair of MMC-induced damage, confirming results reported previously (Lestini *et al.*, 2010). As discussed above with Hef, loss of NreA is additive with loss of Hjc indicating that these two factors act in different pathways of MMC damage repair.

Significantly, all the results obtained with the strains lacking the entire *nreA* open reading frame were precisely phenocopied with cells carrying the *nreA-C∆10* allele encoding an NreA protein (NreA-C∆10) in which the last 10 amino acids of the protein including the entire PIP motif (GLQTDIFDFA, PIP motif conserved residues underlined) are deleted. This result suggests that NreA function in DNA repair is entirely dependent upon its ability to interact directly with PCNA via this motif. We speculate that this interaction is essential either for recruiting NreA to sites of DNA damage or for localising the NreA protein at the replication forks, ready to contribute to the repair of DNA damage encountered by the moving fork such as MMC-induced inter-strand crosslinks.

Finally, It should be noted that for technical reasons we have been unable to demonstrate a direct physical interaction between NreA and PCNA and to confirm that the predicted interaction is indeed dependent upon the presence of the putative PIP motif. Attempts at expressing various tagged and untagged Nre proteins from diverse archaeal species *(Hfx. volcanii, S, solfataricus, H. butylicus, A. fulgidus)* in recombinant form were not successful, nor was it possible to overexpress and purified tagged *S. solfataricus* NreA and NreB proteins from their native host (data not shown). These failures also prevented the search for an enzymatic function for the Nre proteins. As noted in the Introduction, all known archaeal PIP motif-containing PCNA binding proteins are enzymes with DNA-directed activities such as DNA polymerases, nucleases and helicases, which might imply that the Nre proteins have an enzymatic function that is yet to be uncovered. Alternatively, Nre proteins could act as a platform on to which other proteins could assemble or as an inhibitor of interactions between PCNA and the PIP motif-containing enzymes. Further work will be required to resolve these possibilities.

Experimental procedures

Strains and media

Hfx. volcanii strains used in this study are listed in Table 2. All *Hfx. volcanii* strains were grown at 45°C in either rich medium (Hv-YPC) or casamino-acid medium (Hv-CA) prepared as described in the Halohandbook v7.2 (www.haloarchaea.com/resources/halohandbook). For selection procedures, tryptophan (50 µg/ml) and/or thymidine and hypoxanthine (40 µg/ml each) were added to Hv-CA. For counter-selection with 5-fluoroorotic acid (5-FOA), Hv-CA was supplemented with 5-fluoroorotic acid (50 µg/ml) and uracil (10 µg/ml).
Enzymes for molecular cloning were purchased from New England Biolabs (NEB), Promega or Fermentas. Oligonucleotides (Table 3) were synthesized by Integrated DNA Technologies. DNA sequencing was performed by DNA Sequencing & Services, University of Dundee. DNA purification kits were from Qiagen. Diagnostic PCR was performed using MyTaqRed (Bioline) and preparative PCR using Q5 DNA polymerase (NEB). For routine molecular cloning, *E. coli* DH5α (*fhuA2 Δ(argF-lacZ)U169 phoA glnV44 Φ80 Δ(lacZ)M15 gyrA96 recA1 relA1 endA1 thi-1 hsdR17*) was used. To prepare unmethylated plasmid DNA for *Hfx. volcanii* transformation, *E. coli* strain SCS110 (*rpsL* (Str^r) *thr leu endA thi-1 lacY galK galT ara tonA tsx dam dcm supE44* Δ(*lac-proAB*) [F' *traD36 proAB lacI*q*ZΔM15*]) (Stratagene) was used. *E. coli* strains were grown in LB medium (Formedium) supplemented with ampicillin (100 µg/ml) as required.

Plasmids for gene and PIP motif deletion

Plasmids for gene deletion were constructed by separately amplifying the 5' and 3' flanking regions of the *nreA* gene using primer pairs HfxNreA-5S and HfxNreA-5B (5' flanking region) and HfxNreA-3B and HfxNreA-3E (3' flanking region) and *Hfx. volcanii* DS70 genomic DNA as template. Primer sequences are shown in Table 3. The PCR products were then digested using *Spe*I and *Bam*HI (5' flanking region product) or *Bam*HI and *Eco*RI (3' flanking region product fragment), and ligated together in pTA131 (Allers *et al.*, 2010) that had been digested with *Spe*I and *Eco*RI. The resulting plasmid, pTA131-HfxNreA-SBE, was used for markerless deletion of the *nreA* gene. For replacement of *nreA* with the *trpA* selectable marker, a DNA fragment containing the *trp*A selectable marker gene was obtained by restriction of plasmid pTA298 (Allers *et al.*, 2010) with *Bam*HI, and cloned into pTA131-HfxNreA-SBE digested with the same enzyme to make plasmid pTA131-HfxNreA-SBE-TrpA.
The plasmid for PIP motif deletion was constructed by PCR overlap extension mutagenesis (OEM). Initially, two separate PCR reactions were performed with primers HfxNreA-PIP-5E and -PIP-R (first reaction) and HfxNreA-PIP-F and -PIP-3S

(second reaction) and *Hfx. volcanii* DS70 genomic DNA as template. The products of these reactions were partially purified, mixed and used as a template for amplification with primers HfxNreA-PIP-5E and -PIP-3S. The resulting product was restricted with *Eco*RI and *Spe*I and cloned into pTA131 (Allers *et al.*, 2010) cut with the same enzymes, to produce plasmid pTA131-HfxNreA-C∆10.
Plasmids were sequenced to ensure the absence of unwanted changes in the sequence and subsequently transformed into *E. coli* SCS110 to obtain unmethylated DNA for *Hfx. volcanii* transformation.

Gene deletion

The pop-in/pop-out method was used for gene deletion (Bitan-Banin *et al.*, 2003). Unmethylated plasmids pTA131-HfxNreA-SBE-TrpA and pTA131-HfxNreA-SBE were transformed into various *Hfx. volcanii* strains (see Table 2, where plasmid names are shortened to pTA131-NreA-TrpA and pTA131-NreA for brevity) and transformants obtained following growth for 7 days at 45°C on Hv-CA medium supplemented with tryptophan and/or thymidine and hypoxanthine as required. For *trpA* gene replacement, pTA131-HfxNreA-SBE-TrpA transformants of strain H99 were picked, resuspended in Hv-CA medium and plated on Hv-CA plates containing thymidine and hypoxanthine, uracil and 5-FOA with or without tryptophan. Following 7 days growth at 45°C, 16 colonies were picked from plates lacking tryptophan (i.e. putative gene replacements) and screened by PCR using primers HfxNreA-PCR-F and HfxNreA-PCR-R (Table 3); all sixteen appeared to have lost the *nreA* gene. Candidate ∆*nreA::trpA* strains were then re-streaked to single colonies, re-tested to confirm the ∆*nreA* genotype and stored at -80°C in 20% glycerol.
For each markerless gene deletion, a single pTA131-HfxNreA-SBE transformant (i.e. a pop-in colony) was picked, resuspended in Hv-CA medium and then plated on Hv-CA plates containing uracil and 5-FOA (and supplements as required) at 45°C. Following 7 days growth, colonies were picked and screened by PCR as above. Using this method, markerless ∆*nreA* strains were generally recovered at a frequency ranging from 1 in 8 to 1 in 16 of colonies analysed. As above, candidate ∆*nreA* strains were then re-streaked to single colonies, re-tested to confirm the ∆*nreA* genotype and stored at -80°C.

PIP motif deletion

To delete the NreA PIP motif, unmethylated plasmid pTA131-HfxNreA-CΔ10 was transformed into various *Hfx. volcanii* strains and pop-out strains obtained in the presence of uracil and 5-FOA exactly as described above. Pop-out colonies were then screened using primers HfxNreA-PIP-SF and PIP-SR (these produce a 334 nt PCR product in wild-type cells that is reduced to 304 nt in *nreA-CΔ10* cells) or with primer HfxNreA-PIP-SR paired with either of the 3' mistmatched primers HfxNreA-PIP-WT or HfxNreA-PIP-MUT (see Table 3). The HfxNreA-PIP-SR and -PIP-WT pair will only amplify the wild-type chromosome (i.e. with intact PIP motif coding sequence) whereas the HfxNreA-PIP-SR and -PIP-MUT pair will only amplify the *nreA-CΔ10* chromosome.

DNA damage sensitivity assays

Cells were grown in Hv-YPC (with thymidine and hypoxanthine when needed) to an OD_{650} of 0.3 and then 10-fold serial diluted. Dilutions were either plated (100 µl) or spotted (5 µl) on Hv-YPC plates containing increasing amounts of mitomycin C (MMC) ranging from 0 to 40 ng/ml or MMS ranging from 0 to 0.12%. After drying, plates were incubated at 45°C for 5 days. Colonies were then counted and survival curve plotted. Alternatively, serial dilutions were spotted or plated on Hv-YPC plates and exposed to UV-C (254 nm) doses ranging from 0 to 300 J/m^2 using a Stratalinker UV1800 (Agilent) and grown in the dark at 45°C for 5-7 days.

Pulse Field Gel Electrophoresis (PFGE) assays

PFGE was carried out according to the protocol of Kish and co-workers (Kish & DiRuggiero, 2008), with minor modifications. *Hfx. volcanii* cells were grown at 45°C in Hv-YPC to an OD_{650} of 0.3 (t = 0 h) before addition of MMC to a final concentration of 0.5 µg/ml for 1 hour. Cells were harvested by centrifugation at 1800 *g* for 10 minutes, resuspended in pre-warmed Hv-YPC and grown for 10 hours in Hv-YPC to allow recovery. $1x10^9$ cells were sampled at the following times: 0 h, after 1 h of MMC exposure and at 2 h, 4 h, 6 h, 8 h and 10 h (during recovery).
Samples were harvested by centrifugation at 1800 *g* for 10 minutes and resuspended in 500 µl of 18% salt water (SW), prepared as describe in the Halohandbook v7.2 (see above). 500µl of low-melt agarose solution (0.8% agarose in 3:1 18% SW:H_2O) were added, and after gentle mixing, the suspension was poured into plug molds. Cells were then lysed by adding the plugs in proteinase K solution (250 mM EDTA pH8, 1% N-lauryl-sarcosine, 0.5 mg/ml proteinase K) for 18 h at 56°C. Plugs were

washed 2 x 30 minutes in Tris-EDTA (TE) solution (25 mM Tris pH 7.5, 100 mM EDTA) at 37°C. Proteinase K was inactivated overnight by incubating the plugs at 37°C in TE solution containing 0.5 mM PMSF. Plugs were washed twice for 30 minutes in TE solution, once for 30 minutes in 1/10 TE solution and once for 30 minutes in restriction buffer. Each plug was then incubated overnight in 250 µl of restriction buffer with 40 units of *Swa*I enzyme at room temperature.
Prior to electrophoresis, plugs were equilibrated in 0.5X TBE for 1 h. They were then inserted into 1.2% agarose gel (BioRad), and the electrophoresis was run in TBE 0.5X using a CHEF DR III system (BioRad). The running conditions were as follow: 6 V/cm, 0.7 s to 75 s switching time, 120° included angle, at 14°C for 24h.

Acknowledgements

We are grateful to our colleagues, past and present, who contributed to the NreA/MypA project in Copenhagen and St Andrews, including Aleh Razanau, An Zhao, Romy Alisch, Anders Norman and Fiona Gray, to Hugh De Long and Jen Gresham for support, to Roxanne Lestini, Hannu Myllykallio, Thorsten Allers and Jocelyne DiRuggiero for strains and methods, to David Ferrier for use of PFGE equipment and to Malcolm White for valuable suggestions. This work was supported by the USAF Office of Scientific Research under award number FA9550-10-1-0421. The authors declare that they have no conflicts of interests.

References

Allers, T., S. Barak, S. Liddell, K. Wardell & M. Mevarech, (2010) Improved strains and plasmid vectors for conditional overexpression of His-tagged proteins in *Haloferax volcanii*. *Appl Environ Microbiol* **76**: 1759-1769.

Altschul, S.F. & E.V. Koonin, (1998) Iterated profile searches with PSI-BLAST--a tool for discovery in protein databases. *Trends Biochem Sci* **23**: 444-447.

Altschul, S.F., T.L. Madden, A.A. Schaffer, J. Zhang, Z. Zhang, W. Miller & D.J. Lipman, (1997) Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. *Nucleic Acids Res* **25**: 3389-3402.

Bargonetti, J., E. Champeil & M. Tomasz, (2010) Differential toxicity of DNA adducts of mitomycin C. *J Nucleic Acids* **2010**.

Beranek, D.T., (1990) Distribution of methyl and ethyl adducts following alkylation with monofunctional alkylating agents. *Mutat Res* **231**: 11-30.

Bitan-Banin, G., R. Ortenberg & M. Mevarech, (2003) Development of a gene knockout system for the halophilic archaeon *Haloferax volcanii* by use of the pyrE gene. *J Bacteriol* **185**: 772-778.

Boal, A.K., E. Yavin & J.K. Barton, (2007) DNA repair glycosylases with a [4Fe-4S] cluster: a redox cofactor for DNA-mediated charge transport? *Journal of inorganic biochemistry* **101**: 1913-1921.

Brosh, R.M., Jr., (2014) UvrD helicase: the little engine that could. *Cell Cycle* **13**: 1213-1215.

Buis, J.M., J. Cheek, E. Kalliri & J.B. Broderick, (2006) Characterization of an active spore photoproduct lyase, a DNA repair enzyme in the radical S-adenosylmethionine superfamily. *J Biol Chem* **281**: 25994-26003.

Chandor, A., O. Berteau, T. Douki, D. Gasparutto, Y. Sanakis, S. Ollagnier-de-Choudens, M. Atta & M. Fontecave, (2006) Dinucleotide spore photoproduct, a minimal substrate of the DNA repair spore photoproduct lyase enzyme from *Bacillus subtilis*. *J Biol Chem* **281**: 26922-26931.

Crowley, D.J., I. Boubriak, B.R. Berquist, M. Clark, E. Richard, L. Sullivan, S. DasSarma & S. McCready, (2006) The *uvrA*, *uvrB* and *uvrC* genes are required for repair of ultraviolet light induced DNA photoproducts in Halobacterium sp. NRC-1. *Saline Systems* **2**: 11.

De Biasio, A. & F.J. Blanco, (2013) Proliferating cell nuclear antigen structure and interactions: too many partners for one dancer? *Adv Protein Chem Struct Biol* **91**: 1-36.

Delmas, S., I.G. Duggin & T. Allers, (2013) DNA damage induces nucleoid compaction via the Mre11-Rad50 complex in the archaeon *Haloferax volcanii*. *Mol Microbiol* **87**: 168-179.

Delmas, S., L. Shunburne, H.P. Ngo & T. Allers, (2009) Mre11-Rad50 promotes rapid repair of DNA damage in the polyploid archaeon *Haloferax volcanii* by restraining homologous recombination. *PLoS Genet* **5**: e1000552.

Dieckman, L.M., B.D. Freudenthal & M.T. Washington, (2012) PCNA structure and function: insights from structures of PCNA complexes and post-translationally modified PCNA. *Subcell Biochem* **62**: 281-299.

Epshtein, V., (2015) UvrD helicase: an old dog with a new trick: how one step backward leads to many steps forward. *Bioessays* **37**: 12-19.

Epshtein, V., V. Kamarthapu, K. McGary, V. Svetlov, B. Ueberheide, S. Proshkin, A. Mironov & E. Nudler, (2014) UvrD facilitates DNA repair by pulling RNA polymerase backwards. *Nature* **505**: 372-377.

Farkas, J.A., J.W. Picking & T.J. Santangelo, (2013) Genetic techniques for the archaea. *Annu Rev Genet* **47**: 539-561.

Finn, R.D., A. Bateman, J. Clements, P. Coggill, R.Y. Eberhardt, S.R. Eddy, A. Heger, K. Hetherington, L. Holm, J. Mistry, E.L. Sonnhammer, J. Tate & M.

Punta, (2014) Pfam: the protein families database. *Nucleic Acids Res* **42**: D222-230.
Gulbis, J.M., Z. Kelman, J. Hurwitz, M. Odonnell & J. Kuriyan, (1996) Structure of the c-terminal region of $p21^{Waf1/Cip1}$ complexed with human PCNA. *Cell* **87**: 297-306.
Hartman, A.L., C. Norais, J.H. Badger, S. Delmas, S. Haldenby, R. Madupu, J. Robinson, H. Khouri, Q. Ren, T.M. Lowe, J. Maupin-Furlow, M. Pohlschroder, C. Daniels, F. Pfeiffer, T. Allers & J.A. Eisen, (2010) The complete genome sequence of *Haloferax volcanii* DS2, a model archaeon. *PLoS One* **5**: e9605.
Hawkins, M., S. Malla, M.J. Blythe, C.A. Nieduszynski & T. Allers, (2013) Accelerated growth in the absence of DNA replication origins. *Nature* **503**: 544-547.
Hedglin, M., R. Kumar & S.J. Benkovic, (2013) Replication clamps and clamp loaders. *Cold Spring Harb Perpsect Biol* **5**: a010165.
Hering, O., M. Brenneis, J. Beer, B. Suess & J. Soppa, (2009) A novel mechanism for translation initiation operates in haloarchaea. *Mol Microbiol* **71**: 1451-1463.
Kish, A. & J. DiRuggiero, (2008) Rad50 is not essential for the Mre11-dependent repair of DNA double-strand breaks in *Halobacterium* sp. strain NRC-1. *J Bacteriol* **190**: 5210-5216.
Kisker, C., J. Kuper & B. Van Houten, (2013) Prokaryotic nucleotide excision repair. *Cold Spring Harb Perpsect Biol* **5**: a012591.
Kong, X.P., R. Onrust, M. O'Donnell & J. Kuriyan, (1992) Three-dimensional structure of the β subunit of *E. coli* DNA polymerase III holoenzyme: a sliding DNA clamp. *Cell* **69**: 425-437.
Krishna, T.S., X.P. Kong, S. Gary, P.M. Burgers & J. Kuriyan, (1994) Crystal structure of the eukaryotic DNA polymerase processivity factor PCNA. *Cell* **79**: 1233-1243.
Kristensen, T.P., R. Maria Cherian, F.C. Gray & S.A. MacNeill, (2014) The haloarchaeal MCM proteins: bioinformatic analysis and targeted mutagenesis of the β7-β8 and β9-β10 hairpin loops and conserved zinc binding domain cysteines. *Front Microbiol* **5**: 123.
Kuper, J. & C. Kisker, (2013) DNA Helicases in NER, BER, and MMR. *Adv Exp Med Biol* **767**: 203-224.
Lestini, R., Z. Duan & T. Allers, (2010) The archaeal XPF/Mus81/FANCM homolog Hef and the Holliday junction resolvase Hjc define alternative pathways that are essential for cell viability in *Haloferax volcanii*. *DNA Repair (Amst)* **9**: 994-1002.
Lestini, R., S.P. Laptenok, J. Kuhn, M.A. Hink, M.C. Schanne-Klein, U. Liebl & H. Myllykallio, (2013) Intracellular dynamics of archaeal FANCM homologue Hef in response to halted DNA replication. *Nucleic Acids Res* **41**: 10358-10370.
Maga, G. & U. Hubscher, (2003) Proliferating cell nuclear antigen (PCNA): a dancer with many partners. *J Cell Sci* **116**: 3051-3060.
Makarova, K.S., A.V. Sorokin, P.S. Novichkov, Y.I. Wolf & E.V. Koonin, (2007) Clusters of orthologous genes for 41 archaeal genomes and implications for evolutionary genomics of archaea. *Biol Direct* **2**: 33.
Meslet-Cladiere, L., C. Norais, J. Kuhn, J. Briffotaux, J.W. Sloostra, E. Ferrari, U. Hubscher, D. Flament & H. Myllykallio, (2007) A novel proteomic approach identifies new interaction partners for proliferating cell nuclear antigen. *J Mol Biol* **372**: 1137-1148.
Moldovan, G.L., B. Pfander & S. Jentsch, (2007) PCNA, the maestro of the replication fork. *Cell* **129**: 665-679.
Morgunova, E., F.C. Gray, S.A. MacNeill & R. Ladenstein, (2009) Structural insights into the adaptation of proliferating cell nuclear antigen (PCNA) from *Haloferax volcanii* to a high-salt environment. *Acta Crsta* **65**: 1081-1088.
Motz, M., I. Kober, C. Girardot, E. Loeser, U. Bauer, M. Albers, G. Moeckel, E. Minch, H. Voss, C. Kilger & M. Koegl, (2002) Elucidation of an archaeal

replication protein network to generate enhanced PCR enzymes. *J Biol Chem* **277**: 16179-16188.
Palmer, J.R. & C.J. Daniels, (1995) In vivo definition of an archaeal promoter. *J Bacteriol* **177**: 1844-1849.
Poidevin, L. & S.A. MacNeill, (2006) Biochemical characterisation of LigN, an NAD+-dependent DNA ligase from the halophilic euryarchaeon *Haloferax volcanii* that displays maximal in vitro activity at high salt concentrations. *BMC Mol Biol* **7**: 44.
Rebeil, R. & W.L. Nicholson, (2001) The subunit structure and catalytic mechanism of the Bacillus subtilis DNA repair enzyme spore photoproduct lyase. *Proc Natl Acad Sci U S A* **98**: 9038-9043.
Rebeil, R., Y. Sun, L. Chooback, M. Pedraza-Reyes, C. Kinsland, T.P. Begley & W.L. Nicholson, (1998) Spore photoproduct lyase from *Bacillus subtilis* spores is a novel iron-sulfur DNA repair enzyme which shares features with proteins such as class III anaerobic ribonucleotide reductases and pyruvate-formate lyases. *J Bacteriol* **180**: 4879-4885.
Rudolf, J., V. Makrantoni, W.J. Ingledew, M.J. Stark & M.F. White, (2006) The DNA repair helicases XPD and FancJ have essential iron-sulfur domains. *Mol Cell* **23**: 801-808.
Scharer, O.D., (2013) Nucleotide excision repair in eukaryotes. *Cold Spring Harb Perspect Biol* **5**: a012609.
Skowyra, A. & S.A. MacNeill, (2012) Identification of essential and non-essential single-stranded DNA-binding proteins in a model archaeal organism. *Nucleic Acids Res* **40**: 1077-1090.
Slieman, T.A., R. Rebeil & W.L. Nicholson, (2000) Spore photoproduct (SP) lyase from Bacillus subtilis specifically binds to and cleaves SP (5-thyminyl-5,6-dihydrothymine) but not cyclobutane pyrimidine dimers in UV-irradiated DNA. *J Bacteriol* **182**: 6412-6417.
Stroud, A., S. Liddell & T. Allers, (2012) Genetic and Biochemical Identification of a Novel Single-Stranded DNA-Binding Complex in *Haloferax volcanii*. *Front Microbiol* **3**: 224.
Tatusov, R.L., N.D. Fedorova, J.D. Jackson, A.R. Jacobs, B. Kiryutin, E.V. Koonin, D.M. Krylov, R. Mazumder, S.L. Mekhedov, A.N. Nikolskaya, B.S. Rao, S. Smirnov, A.V. Sverdlov, S. Vasudevan, Y.I. Wolf, J.J. Yin & D.A. Natale, (2003) The COG database: an updated version includes eukaryotes. *BMC Bioinformatics* **4**: 41.
Tatusov, R.L., M.Y. Galperin, D.A. Natale & E.V. Koonin, (2000) The COG database: a tool for genome-scale analysis of protein functions and evolution. *Nucleic Acids Res* **28**: 33-36.
Tatusov, R.L., D.A. Natale, I.V. Garkavtsev, T.A. Tatusova, U.T. Shankavaram, B.S. Rao, B. Kiryutin, M.Y. Galperin, N.D. Fedorova & E.V. Koonin, (2001) The COG database: new developments in phylogenetic classification of proteins from complete genomes. *Nucleic Acids Res* **29**: 22-28.
Tittel-Elmer, M., C. Alabert, P. Pasero & J.A. Cobb, (2009) The MRX complex stabilizes the replisome independently of the S phase checkpoint during replication stress. *EMBO J* **28**: 1142-1156.
Tomasz, M., (1995) Mitomycin C: small, fast and deadly (but very selective). *Chem Biol* **2**: 575-579.
Tsurimoto, T., (2006) PCNA-interacting proteins. In: Proliferating cell nuclear antigen (PCNA). H. Lee (ed). Kerala: Research Signpost, pp. 25-49.
Warbrick, E., (1998) PCNA binding through a conserved motif. *Bioessays* **20**: 195-199.
Wendoloski, D., C. Ferrer & M.L. Dyall-Smith, (2001) A new simvastatin (mevinolin)-resistance marker from *Haloarcula hispanica* and a new *Haloferax volcanii* strain cured of plasmid pHV2. *Microbiology* **147**: 959-964.

Winter, J.A. & K.A. Bunting, (2012) Rings in the extreme: PCNA interactions and adaptations in the archaea. *Archaea (Vancouver, B.C* **2012**: 951010.
Winter, J.A., P. Christofi, S. Morroll & K.A. Bunting, (2009) The crystal structure of *Haloferax volcanii* proliferating cell nuclear antigen reveals unique surface charge characteristics due to halophilic adaptation. *BMC Struct Biol* **9**: 55.
Wolf, Y.I., K.S. Makarova, N. Yutin & E.V. Koonin, (2012) Updated clusters of orthologous genes for Archaea: a complex ancestor of the Archaea and the byways of horizontal gene transfer. *Biol Direct* **7**: 46.
Yao, N.Y. & M. O'Donnell, (2012) The RFC clamp loader: structure and function. *Subcell Biochem* **62**: 259-279.
Zhao, A., F.C. Gray & S.A. MacNeill, (2006) ATP- and NAD^+-dependent DNA ligases share an essential function in the halophilic archaeon *Haloferax volcanii*. *Mol Microbiol* **59**: 743-752.

Table 1: Nre proteins in diverse archaeal lineages				
		NreA	**NreB**	**NreB-linked SplB**
Euryarchaeota				
Thermococcales	*Thermococcus kodakaraensis*	+		
	Thermococcus sibiricus	+	+	
	Pyrococcus furiosus	+		
Methanopyrales	*Methanopyrus kandleri*	+		
Methanobacteriales	*Methanosphaera stadtmanae*	+		
Methanococcales	*Methanococcus maripaludis*	+		
Methanocella	*Methanocella paludicola*	+	+	+
Thermoplasmatales	*Picrophilus torridus*		+	+
Archaeoglobales	*Archaeoglobus fulgidus*	+	+	+
Methanosarcinales	*Methanosarcina acetivorans*	+		
Methanomicrobiales	*Methanospirillum hungatei*	+		
Halobacteriales	*Haloferax volcanii*	+		
Thaumarchaeota				
	Cenarchaeum symbiosum	+		
Aigarchaeota				
	Caldiarchaeum subterraneum	+		
Crenarchaeota				
Thermoproteales	*Thermofilum pendens*	+	+	+
	Pyrobaculum aerophilum	+	+	
Sulfolobales	*Sulfolobus solfataricus*	+	+	+
Desulfurococcales	*Hyperthermus butylicus*	+		
	Staphylothermus marinus		+	+
Korarchaeota				
	Korarchaeum cryptofilum	+	+	

Table 2: *Hfx. volcanii* strains used in this study		
Strain no.	Genotype	Source and reference, or method of construction
SMH693		DS70 (Wendoloski *et al.*, 2001)
SMH630	Δ*pyrE2*	H26 (Allers *et al.*, 2010)
SMH628	Δ*trpA* Δ*pyrE2*	H53 (Allers *et al.*, 2010)
SMH892	Δ*hdrB* Δ*pyrE2*	H98 (Allers *et al.*, 2010)
SMH826	Δ*trpA* Δ*hdrB* Δ*pyrE2*	H99 (Allers *et al.*, 2010)
SMH894	Δ*nreA::trpA* Δ*trpA* Δ*hdrB* Δ*pyrE2*	SMH826, pTA131-NreA-TrpA
SMH795	*ΔpyrE2 bgaHa-Kp*	H115 (Delmas *et al.*, 2009)
SMH794	*ΔpyrE2 bgaHa-Kp ΔnreA*	SMH795, pTA131-NreA
SMH859	*ΔpyrE2 bgaHa-Kp nreA-CΔ10*	SMH795, pTA131-NreA-CΔ10
SMH797	*ΔpyrE2 bgaHa-Kp Δmre11 Δrad50*	H204 (Delmas *et al.*, 2009)
SMH796	*ΔpyrE2 bgaHa-Kp Δmre11 Δrad50 ΔnreA*	SMH797, pTA131-NreA
SMH861	*ΔpyrE2 bgaHa-Kp Δmre11 Δrad50 nreA-CΔ10*	SMH797, pTA131-NreA-CΔ10
SMH815	*ΔpyrE2 ΔtrpA ΔhdrB Δhjc*	H282 (Lestini *et al.*, 2010)
SMH835	*ΔpyrE2 ΔtrpA ΔhdrB Δhjc ΔnreA::trpA*	SMH815, pTA131-NreA-TrpA
SMH863	*ΔpyrE2 ΔtrpA ΔhdrB Δhjc nreA-CΔ10*	SMH815, pTA131-NreA-CΔ10
SMH816	*ΔpyrE2 ΔtrpA ΔhdrB Δhef*	H364 (Lestini *et al.*, 2010)
SMH836	*ΔpyrE2 ΔtrpA ΔhdrB Δhef ΔnreA::trpA*	SMH816, pTA131-NreA-TrpA
SMH864	*ΔpyrE2 ΔtrpA Δhdrb Δhef nreA-CΔ10*	SMH816, pTA131-NreA-CΔ10
SMH817	*ΔpyrE2 ΔuvrA*	H509 (Lestini *et al.*, 2010)
SMH837	*ΔpyrE2 ΔuvrA ΔnreA*	SMH817, pTA131-NreA
SMH865	*ΔpyrE2 ΔuvrA nreA-CΔ10*	SMH817, pTA131-NreA-CΔ10
SMH849	*ΔpyrE2 ΔuvrB*	H1181 (Lestini *et al.*, 2010)
SMH869	*ΔpyrE2 ΔuvrB ΔnreA*	SMH849, pTA131-NreA
SMH870	*ΔpyrE2 ΔuvrB nreA-CΔ10*	SMH849, pTA131-NreA-CΔ10
SMH823	*ΔpyrE2 ΔuvrC*	H1187 (Lestini *et al.*, 2010)
SMH850	*ΔpyrE2 ΔuvrC ΔnreA*	SMH823, pTA131-NreA
SMH872	*ΔpyrE2 ΔuvrC nreA-CΔ10*	SMH823, pTA131-NreA-CΔ10
SMH819	*ΔpyrE2 ΔuvrD*	H514 (Lestini *et al.*, 2010)
SMH839	*ΔpyrE2 ΔuvrD ΔnreA*	SMH819, pTA131-NreA
SMH867	*ΔpyrE2 ΔuvrD nreA-CΔ10*	SMH819, pTA131-NreA-CΔ10

Table 3: Oligonucleotide primers used in this study

Restriction sites in primers used for plasmid construction are underlined. Oligonucleotide primers used for mutant detection are shown with 3' mismatched based underlined (HfxNreA-PIP-SF and -SR).

A. For construction of pTA131-HfxNreA-SBE	
HfxNreA-5S	5'-ACTAGTAGGACGACGAGACATCCGAAGAAG-3'
HfxNreA-5B	5'-GGATCCTCACCGGACGTGGGTCGCCCGCGG-3'
HfxNreA-3B	5'-GGATCCGCAACTATTTTGACAGTCGGATAA-3'
HfxNreA-3E	5'-GAATTCTCTACGCGTTTCACCACGTCGGCA-3'
B. For construction of pTA131-HfxNreA-CΔ10	
HfxNreA-PIP-5E	5'-AGTGAATTCTCGAAATCTACCGCAACGAGTTCAT-3'
HfxNreA-PIP-F	5'-GATGGCCGCGTGAGCAACTATTTTGACAGTCGGAT-3'
HfxNreA-PIP-R	5'-TAGTTGCTCACGCGGCCATCGTGGACTTC-3'
HfxNreA-PIP-3S	5'-AGTACTAGTGGTCGTCTACCACCTGAGCATCAT-3'
C. For detection of deletion strains by PCR	
HfxNreA-PCR-F	5'-TTCCCTCGAAGAGATTTCCG-3'
HfxNreA-PCR-R	5'-GTGGTCATCTGGGAGGAAG-3'
D. For detection of PIP motif deletion strains by PCR	
HfxNreA-PIP-WT	5'-TCCACGATGGCCGCGGGC-3'
HfxNreA-PIP-MUT	5'-TCCACGATGGCCGCGTGAG-3'
HfxNreA-PIP-SF	5'-CAGACACGTCTCCGACGACTACT-3'
HfxNreA-PIP-SR	5'-CGATTACGGGGAGGAGCGCGAAA-3'

Figure legends

Fig. 1. A. Schematic representation of thirteen Nre proteins from nine diverse archaeal species showing the presence or absence of the N-terminal C4 metal binding domain (gray box) and C-terminal PCNA binding motif (PIP motif, hatched box). Protein ID and UniProtKB accession numbers: Crenarchaeal species – *Sulfolobus solfataricus* Sso NreA (Sso0872, Q97ZM0), *S. solfataricus* Sso NreB (Sso2743, Q97V92), *Pyrobaculum aerophilum* Pae NreA (PAE1269, Q8ZXI6), *P. aerophilum* Pae NreB (PAE2419, Q8ZV76); Euryarchaeal species – *Archaeoglobus fulgidus* Afu NreA (AF1346-AF1347 fusion, see Experimental procedures for further details), *A. fulgidus* Afu NreB (AF1590, O28682), *Thermococcus kodakaraensis* Tko NreA (TKO1067, Q5JE14), *Pyrococcus abyssi* Pab NreA (PAB2062, Q9V1K0), *Methanococcus maripaludis* Mma NreA (MmarC5_0593, A4FXH8), *Haloferax volcanii* Hvo NreA (HVO_0734, D4GTA5); Thaumarchaeal (Th.) species – *Cenarchaeum symbiosum* Csy NreA (CENSYa_0438, A0RUQ6); Korarchaeal (Kor.) species – *Korarchaeum cryptofilum* Kcr NreA (Kcr_0314, B1L3P3), *K. cryptofilum* Kcr NreB (Kcr_1144, B1L611). **B.** Multiple sequence alignment of C-terminal PCNA binding motif (PIP motif). Conserved residues are boxed. **C.** Multiple amino acid sequence alignment of predicted N-terminal metal binding domain. Conserved cysteines are boxed. **D.** Chromosome context of Nre coding genes in selected species. Additional protein ID and UniProtKB accession numbers: *S. solfataricus* SplB-like protein Sso2744 (Q97V91), *Picrophilus torridus* NreB (PTO1501, Q6KYW6) and SplB-like proteins (PTO1500, Q6KYW7).

Fig. 2. A. Chromosome context of the *Hfx. volcanii HVO_0734* gene encoding NreA. **B.** Multiple sequence alignment of PCNA binding motifs (PIP motifs) in eight *Hfx. volcanii* DNA replication and repair factors, including DNA polymerases B, D and Y (UniProt accession numbers D4GUK5, D4GYR0 and D4GXE6 respectively), the nuclease Fen1 (D4GXI8), RNase HII subunit RnhB (D4GTK7), uracil DNA glycosylase Udg3 (D4GRY5), DNA helicase Hel308b (D4GYK9) and the large subunit of replication factor C, RfcB (D4GSN1). **C.** Growth of *Hfx. volcanii* lacking NreA or expressing NreA-C∆10. Wild-type, ∆*nreA* and *nreA-C∆10* strains were grown to mid-log phase in Hv-YPC medium before being serially diluted (tenfold at each step) in 18% SW and spotted onto Hv-YPC plates. The plates were then incubated for 4 days at 45°C. **D.** ∆*nreA* and *nreA-C∆10* strains are sensitive to MMC. Cells were grown as above and serial dilutions plated on Hv-YPC plates containing 0, 20 or 40 ng/ml MMC. Plates were incubated at 45°C until colonies formed.

Fig. 3: PFGE analysis of *Hfx. volcanii* chromosomal DNA following MMC treatment. Exponentially growing cells were exposed to 0.5 µg/ml MMC for 1 hour and the repair of the chromosome followed by PFGE over a period of 10 hours. Left hand panel: wild-type cells. Centre panel: Δ*nreA* cells. Right hand panel: *nreA-CΔ10* cells. Prior to PFGE, chromosome plugs were digested with *Swa*I restriction enzyme. This enzyme cuts the main chromosome of *Hfx. volcanii* lab strains in five places, producing fragments of 1.7 Mb, 1.43 Mb, 230 kb, 82 kb and 26 kb (Hawkins *et al.*, 2013, Hartman *et al.*, 2010). *Swa*I also cuts the endogenous plasmid pHV3 at two locations, producing 438 kb and 20 kb fragments. Neither the 26 kb (main chromosome) fragment, the 20 kb (pHV3) fragment, nor pHV1, is visible on the PFGE gels.

Fig. 4. Epistasis analysis with known DNA repair factors. **A – G.** Wild-type and Δ*nreA* cells (in various genetic backgrounds) were plated on Hv-YPC plates containing increasing concentrations (0 - 40 ng/ml) of MMC and incubated at 45°C. Survival curves were plotted following growth for 5-7 days. **H.** Wild-type, Δ*nreA*, Δ*uvrA* and Δ*nreA* Δ*uvrA* cells were plated on Hv-YPC plates, exposed to varying does of UV (0 - 300 J/m^2) and survival curves plotted as above. Full strain genotypes are shown in Table 2.

A novel archaeal DNA repair factor that acts with the UvrABC system to repair mitomycin C-induced DNA damage in a PCNA-dependent manner

Xavier Giroux and Stuart A. MacNeill

Supplementary information

In the Pfam database (Finn *et al.*, 2014), N- and C-terminal regions of the Nre proteins are designated as DUF650 (domain of unknown function 650) and DUF651. This designation (specifically, the splitting of the Nre protein into two separate domains) is based on protein sequences encoded by the genome of the *A. fulgidus*: in this organism, sequences encoding the N-terminal and C-terminal domains were reported (Klenk *et al.*, 1997) to be present as separate but adjacent ORFs encoding proteins AF1346 and AF1347, with the C-terminal end of AF1347 being identified in the earlier two-hybrid screen as binding PCNA (Motz *et al.*, 2002). However, the sequence of this part of *A. fulgidus* chromosome is incorrect and both domains are actually expressed as a single AF1346-AF1347 ORF

Resequencing of AF1346-AF1347

As part of an effort to co-express the *A. fulgidus* AF1346 and AF1347 proteins in recombinant form in *E.coli*, the AF1346-AF1347 genomic region was amplified by PCR using primers AfuNreA-5Bam (5'-GTGTGGTTGTGGATCCGATGGTTTCTGTGAGGTGCGCTGAA-3') and AfuNreA-3Pac (5'-GTGTGTGTGTTTAATTAACTTTCCCTCGATTTCCTAACC-3'), restricted with *Bam*HI and *Pac*I (sites underlined in primer sequences) and cloned into the *E.coli* expression vector pRSFDuet-1 (Novagen) that had been cut with the same enzymes. Subsequent sequencing of multiple independent pRSFDuet-1-AfuNreA clones revealed the presence of an additional G after nucleotide 834 of the AF1346 ORF (nucleotide 1204996 in the *A. fulgidus* genome sequence in the GenBank database, accession number AE000782.1), resulting in frameshift and fusion to the AF1347 ORF (F.C. Gray and S.M., unpublished results). The fused ORF is 1173 nucleotides in length and encodes a protein of 391 amino acids that is closely related to NreA proteins encoded by the related species *A. veneficus* (UniProtKB accession

number F2KPG6) and *A. sulfaticallidus* (UniProtKB N0BJ96). A correction to the database entry has been submitted.

Finn, R.D., A. Bateman, J. Clements, P. Coggill, R.Y. Eberhardt, S.R. Eddy, A. Heger, K. Hetherington, L. Holm, J. Mistry, E.L. Sonnhammer, J. Tate & M. Punta, (2014) Pfam: the protein families database. *Nucleic Acids Res* **42**: D222-230.

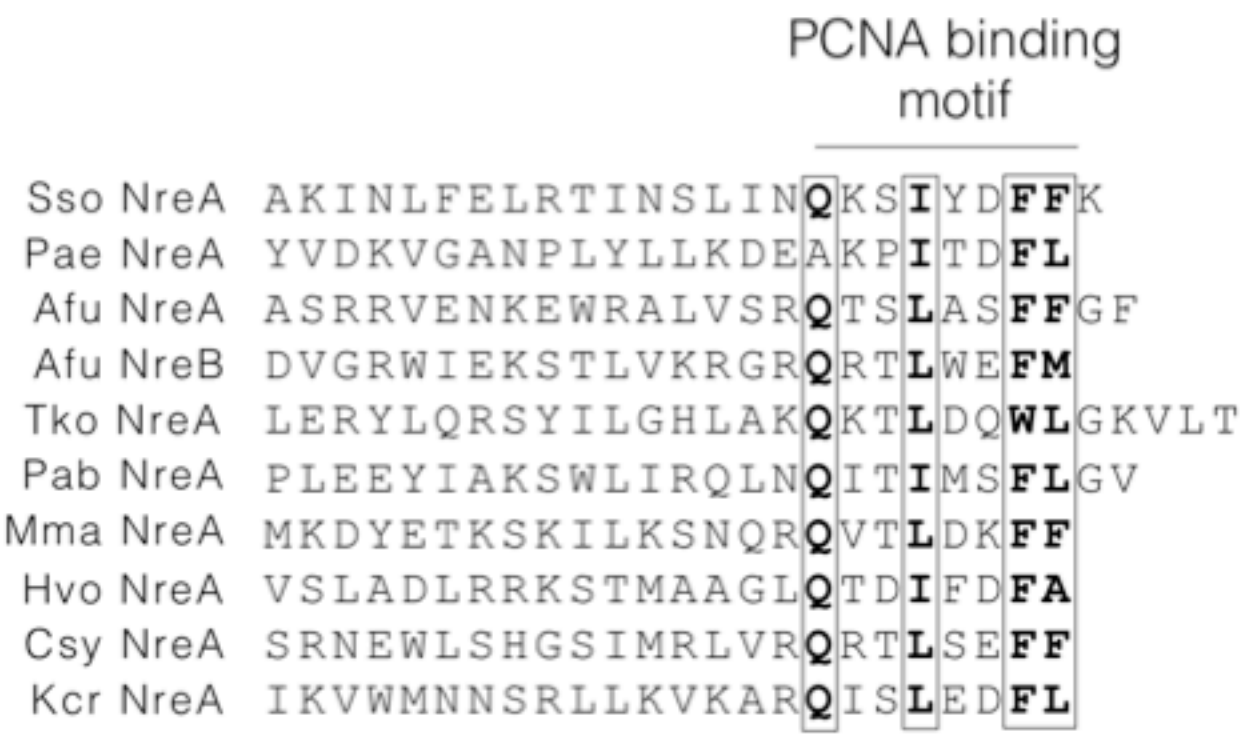
Klenk, H.P., R.A. Clayton, J.F. Tomb, O. White, K.E. Nelson, K.A. Ketchum, R.J. Dodson, M. Gwinn, E.K. Hickey, J.D. Peterson, D.L. Richardson, A.R. Kerlavage, D.E. Graham, N.C. Kyrpides, R.D. Fleischmann, J. Quackenbush, N.H. Lee, G.G. Sutton, S. Gill, E.F. Kirkness, B.A. Dougherty, K. McKenney, M.D. Adams, B. Loftus, J.C. Venter & et al., (1997) The complete genome sequence of the hyperthermophilic, sulphate-reducing archaeon *Archaeoglobus fulgidus*. *Nature* **390**: 364-370.

Motz, M., I. Kober, C. Girardot, E. Loeser, U. Bauer, M. Albers, G. Moeckel, E. Minch, H. Voss, C. Kilger & M. Koegl, (2002) Elucidation of an archaeal replication protein network to generate enhanced PCR enzymes. *J Biol Chem* **277**: 16179-16188.

A


C4 metal binding domain
DUF650–DUF651 domain
PCNA binding motif
N
C
Sso0872
Sso2743
PAE1269
PAE2419
AF1346-AF1347
AF1590
TKO1067
PAB2062
MmarC5_0593
HVO_0734
CENSYa_0438
Kcr_0314
Kcr_1144
Sso NreA
Sso NreB
Pae NreA
Pae NreB
Afu NreA
Afu NreB
Tko NreA
Pab NreA
Mma NreA
Hvo NreA
Csy NreA
Kcr NreA
Kcr NreB
Crenarchaea
Euryarchaea
Th.
Kor.
100 amino acids

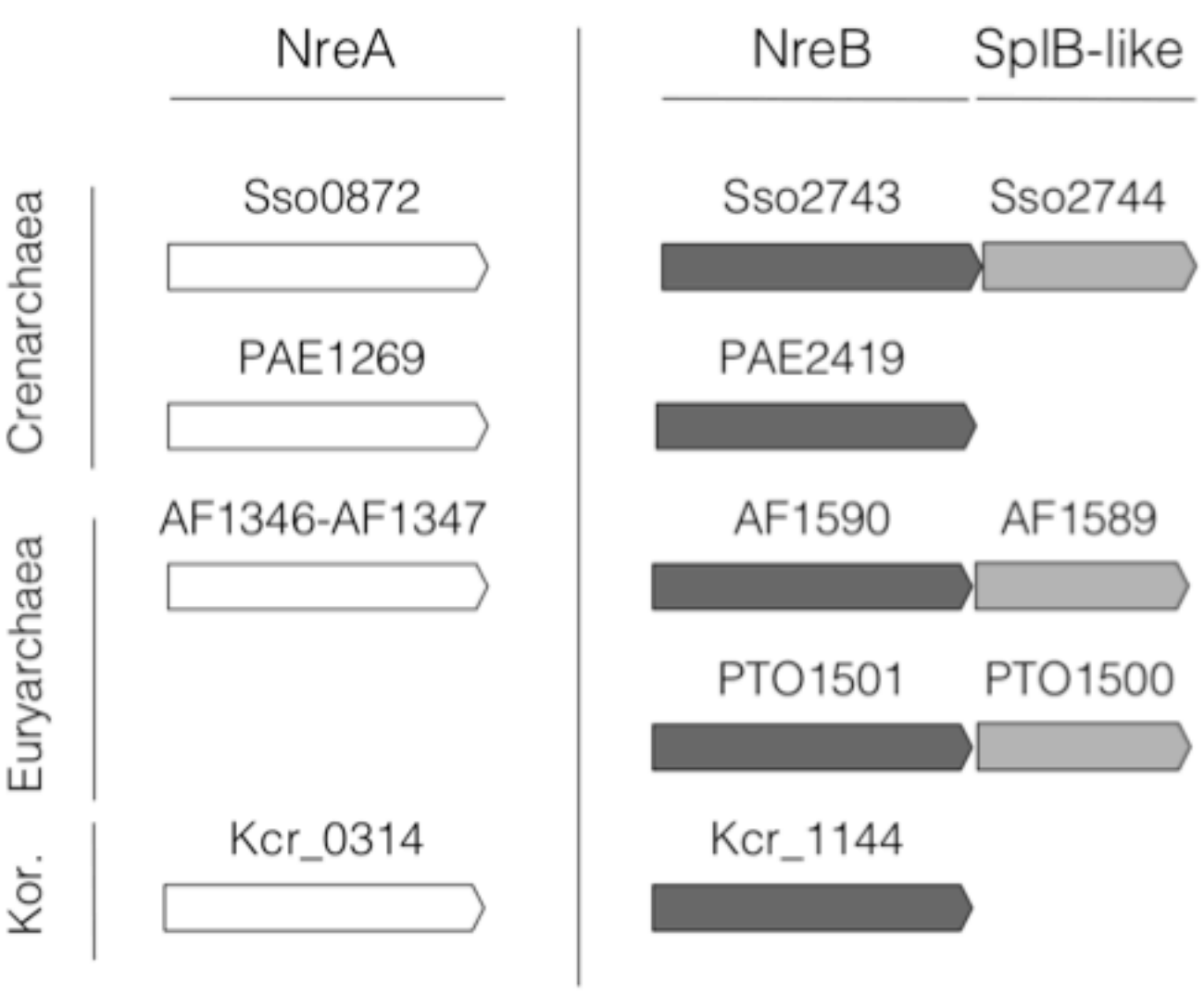
B


PCNA binding motif
Sso NreA AKINLFELRTINSLINQKSIYDFFK
Pae NreA YVDKVGANPLYLLKDEAKPITDFL
Afu NreA ASRRVENKEWRALVSRQTSLASFFGF
Afu NreB DVGRWIEKSTLVKRGRQRTLWEFM
Tko NreA LERYLQRSYILGHLAKQKTLDQWLGKVLT
Pab NreA PLEEYIAKSWLIRQLNQITIMSFLGV
Mma NreA MKDYETKSKILKSNQRQVTLDKFF
Hvo NreA VSLADLRRKSTMAAGLQTDIFDFA
Csy NreA SRNEWLSHGSIMRLVRQRTLSEFF
Kcr NreA IKVWMNNSRLLKVKARQISLEDFL

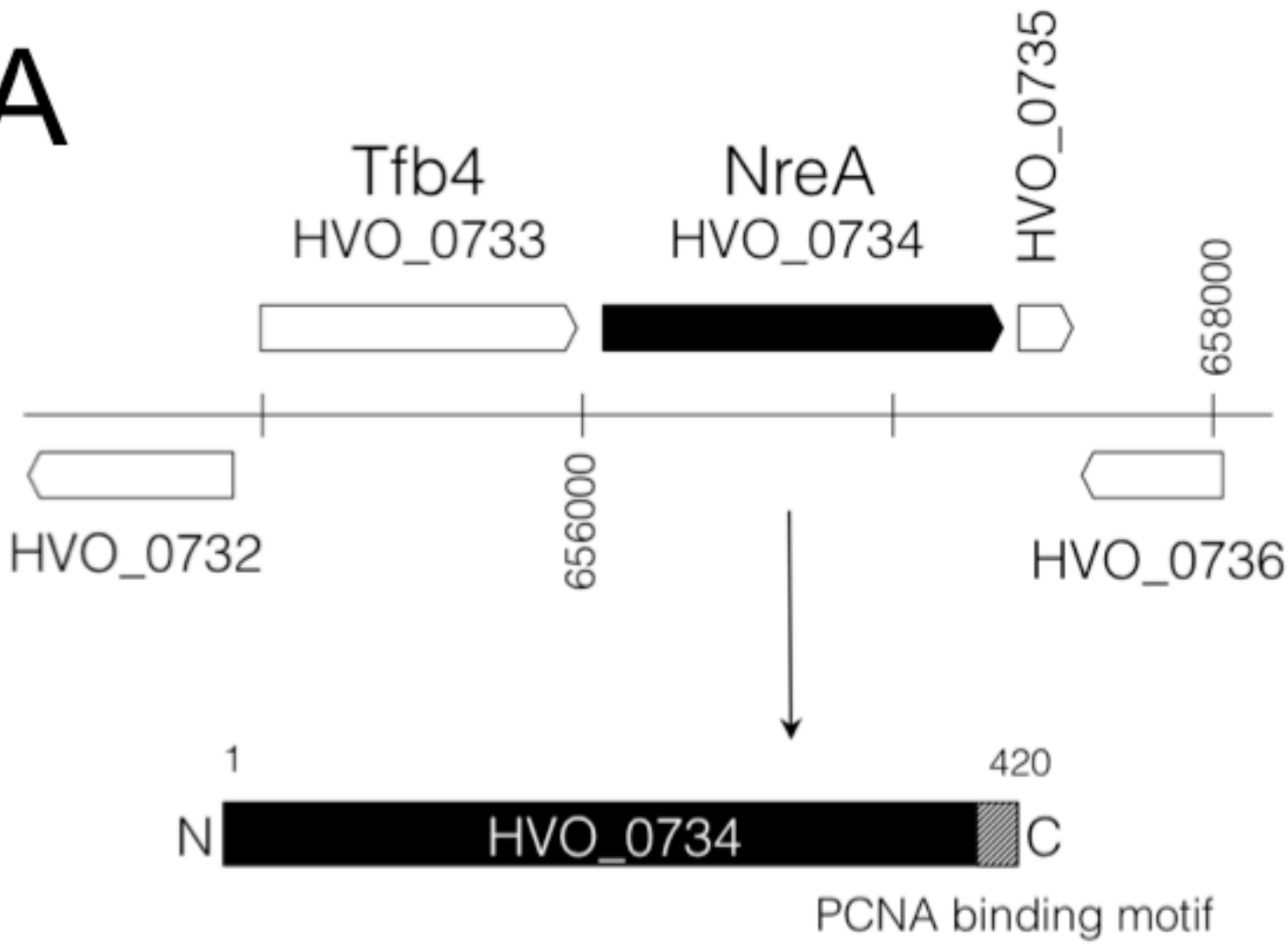
C


C4 metal binding domain
MHKIPAELCIKCKSHKNLCGLPYCPIMERFRG Sso NreA
MIRPELCIRCRGAKYLCGLSYCPIIVNTKT Sso NreB
MSICLKCQVRKLLCGLPRCPIADRISV Pae NreA
MNHSERIGGDLCVKCRGGRYLCGLSYCPLLVRQLS Pae NreB
MVSVRCAECKG-KLLCGRSKCPLLEKYRF Afu NreA
MDCIECRG-RMLCSRKVCPLMPSLRP Afu NreB
MTELFNSKLCAICKGRKLLCGRPTCPILERFRV Tko NreA
MIKVNSELCAICKGRKLLCGRPVCPILERFRI Pab NreA
MELFRSKTCALCKGRKLLCGRPRCPILEKFRV Mma NreA
MSPKINADFCVRCKGSKMLCGLKRCPIVERIKF Kcr NreB
C C C C

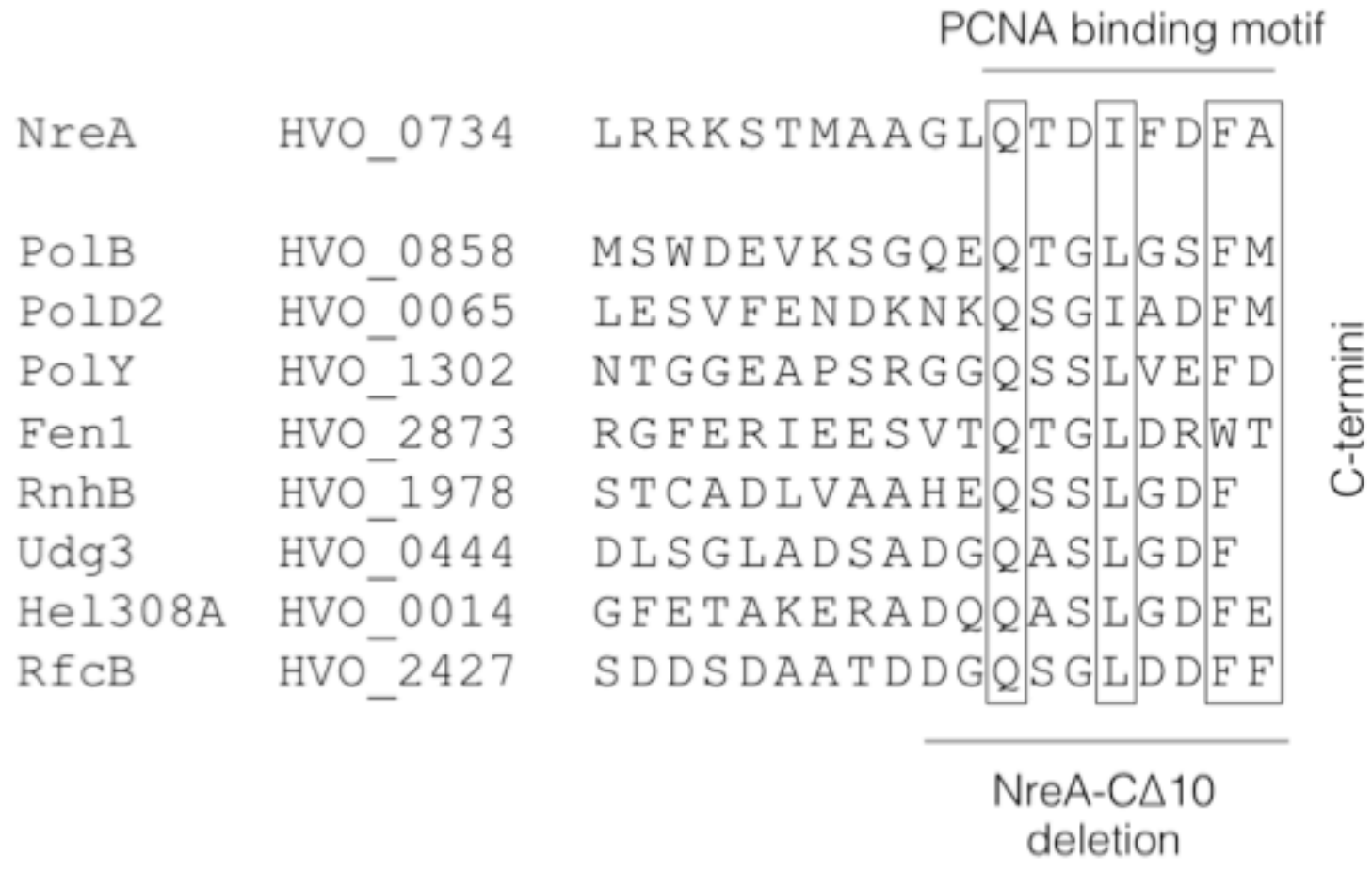
D


NreA
NreB
SplB-like
Crenarchaea
Euryarchaea
Kor.
Sso0872
Sso2743
Sso2744
PAE1269
PAE2419
AF1346-AF1347
AF1590
AF1589
PTO1501
PTO1500
Kcr_0314
Kcr_1144

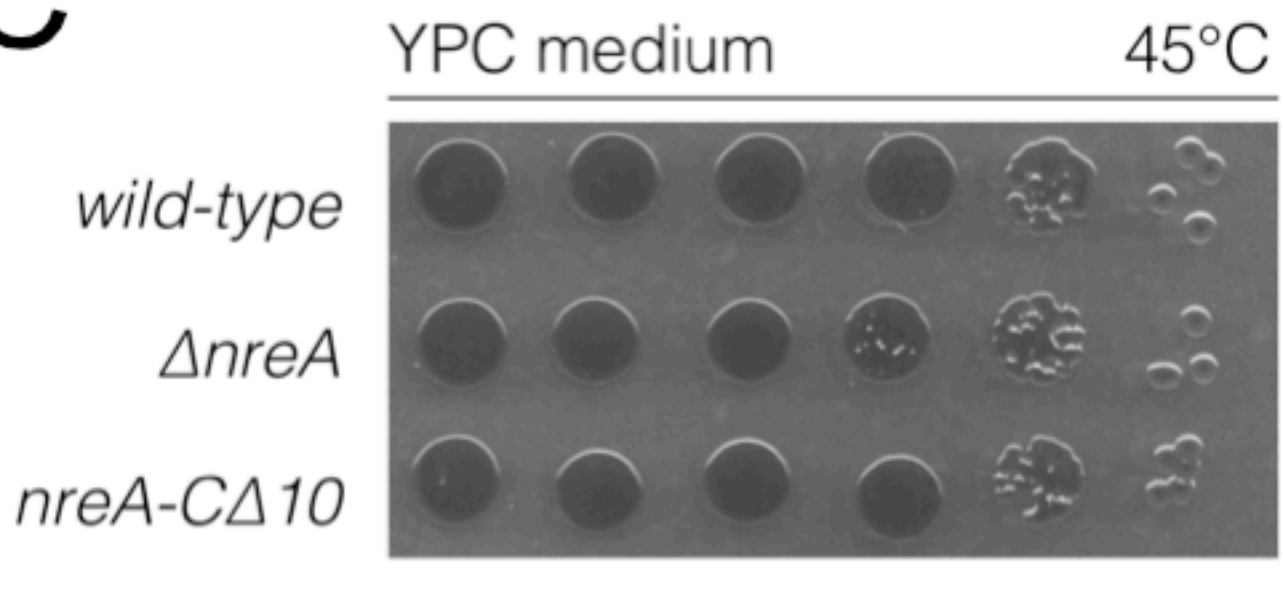
A


Tfb4
HVO_0733
NreA
HVO_0734
HVO_0735
658000
656000
HVO_0732
HVO_0736
1
420
N
HVO_0734
C
PCNA binding motif

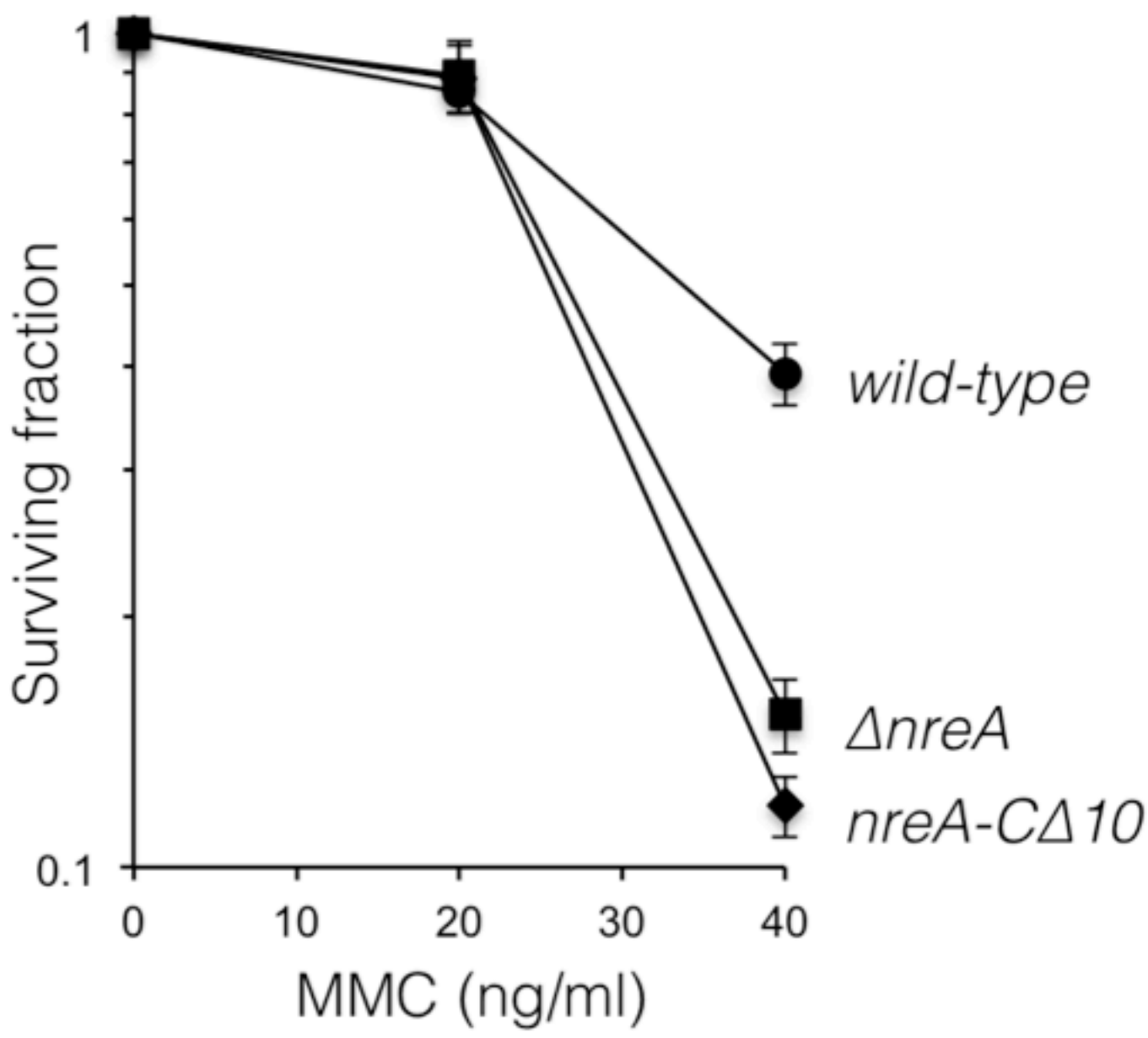
B


PCNA binding motif
NreA HVO_0734 LRRKSTMAAGLQTDIFDFA
PolB HVO_0858 MSWDEVKSGQEQTGLGSFM
PolD2 HVO_0065 LESVFENDKNKQSGIADFM
PolY HVO_1302 NTGGEAPSRGGQSSLVEFD
Fen1 HVO_2873 RGFERIEESVTQTGLDRWT
RnhB HVO_1978 STCADLVAAHEQSSLGDF
Udg3 HVO_0444 DLSGLADSADGQASLGDF
Hel308A HVO_0014 GFETAKERADQQASLGDFE
RfcB HVO_2427 SDDSDAATDDGQSGLDDFF
C-termini
NreA-CΔ10
deletion

C


YPC medium
45°C
wild-type
ΔnreA
nreA-CΔ10

D


1
0.1
Surviving fraction
0
10
20
30
40
MMC (ng/ml)
wild-type
ΔnreA
nreA-CΔ10

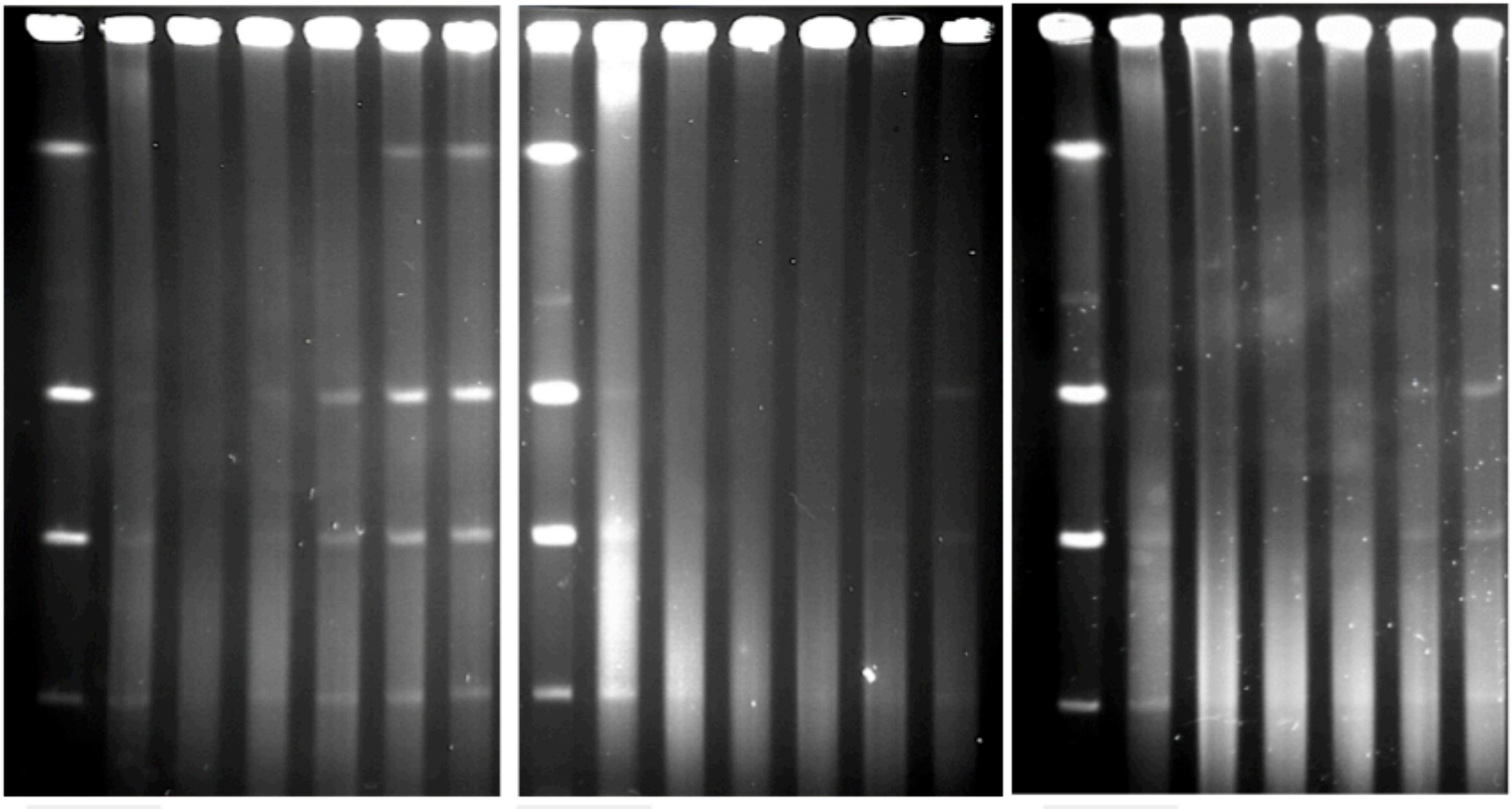
wild-type

ΔnreA

nreA-CΔ10

MMC Recovery MMC Recovery MMC Recovery

kb



1400/1700

418

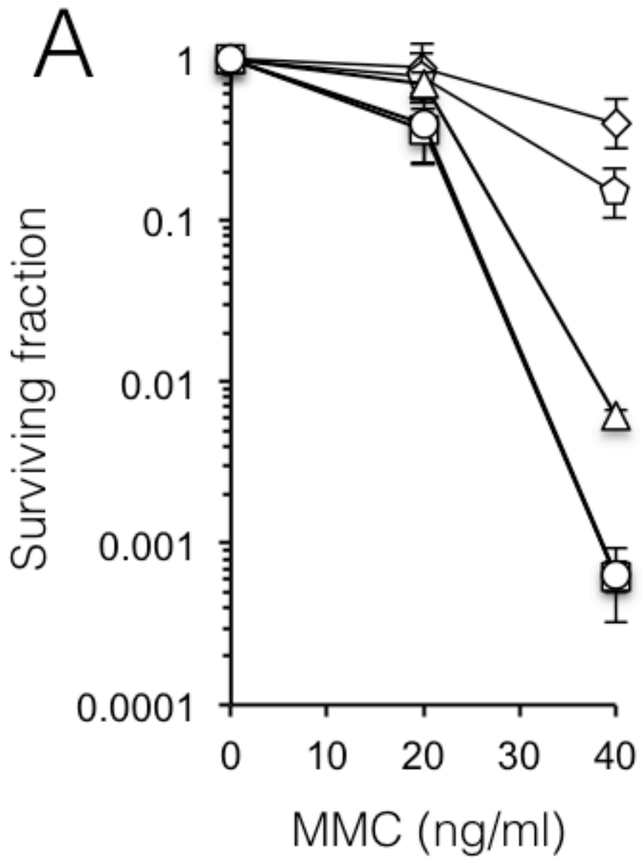
230

82

0 1 2 4 6 8 10 0 1 2 4 6 8 10 0 1 2 4 6 8 10

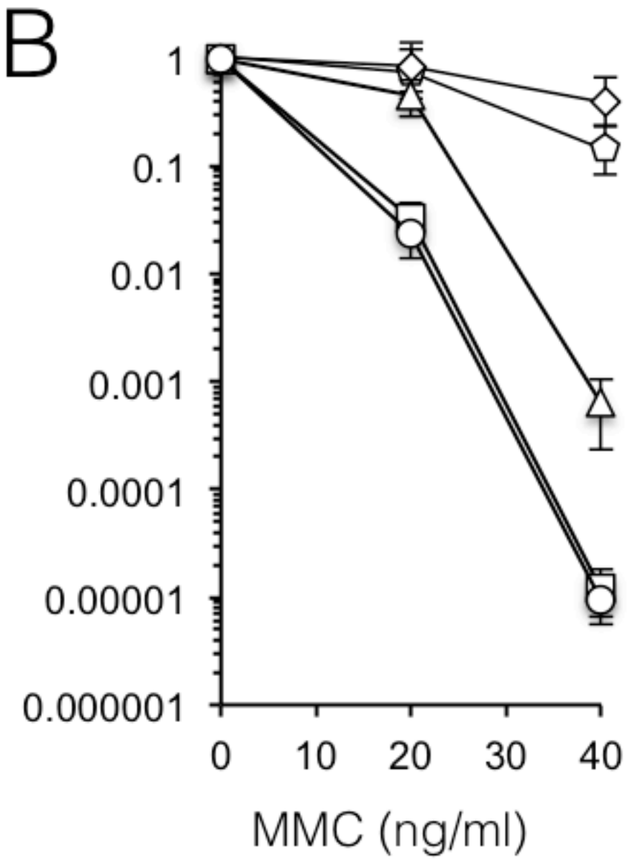
hours hours hours

A



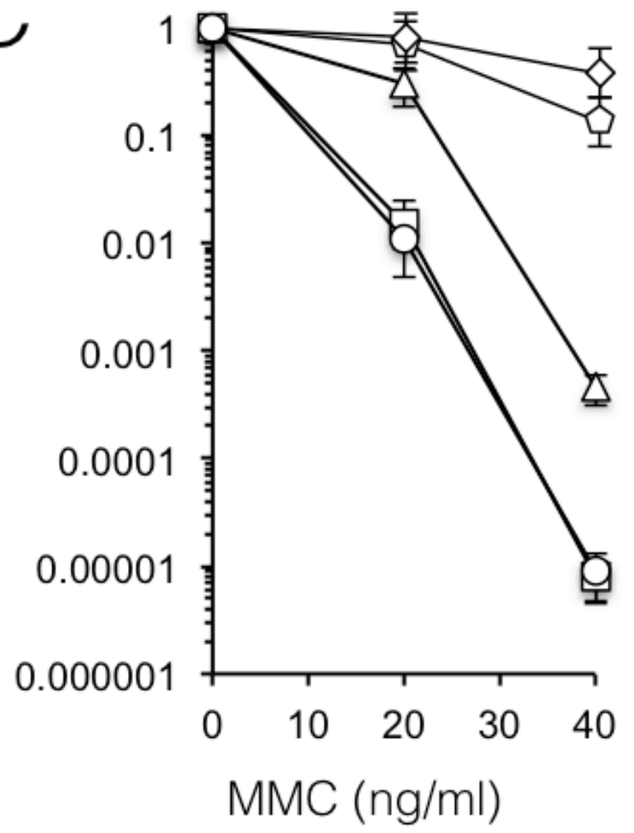
◇ *wild-type*
⬠ *ΔnreA*
△ *Δmre11 Δrad50*
□ *Δmre11 Δrad50 ΔnreA*
○ *Δmre11 Δrad50 nreA-CΔ10*

B



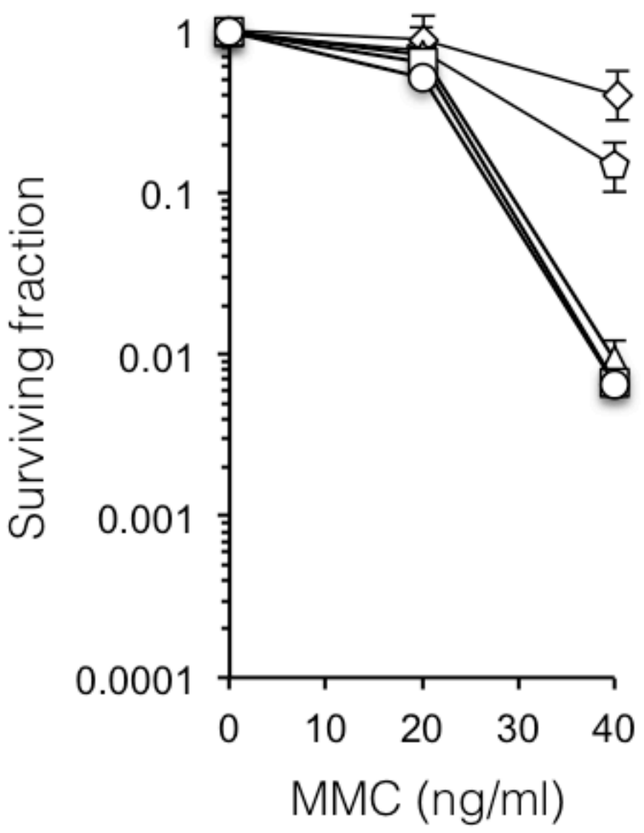
◇ *wild-type*
⬠ *ΔnreA*
△ *Δhjc*
□ *Δhjc ΔnreA*
○ *Δhjc nreA-CΔ10*

C



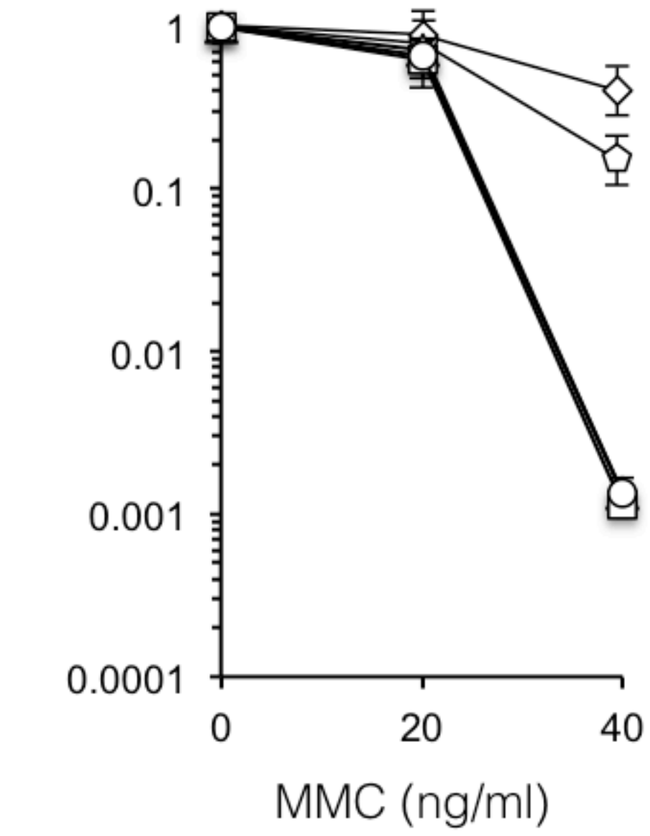
◇ *wild-type*
⬠ *ΔnreA*
△ *Δhef*
□ *Δhef ΔnreA*
○ *Δhef nreA-CΔ10*

D



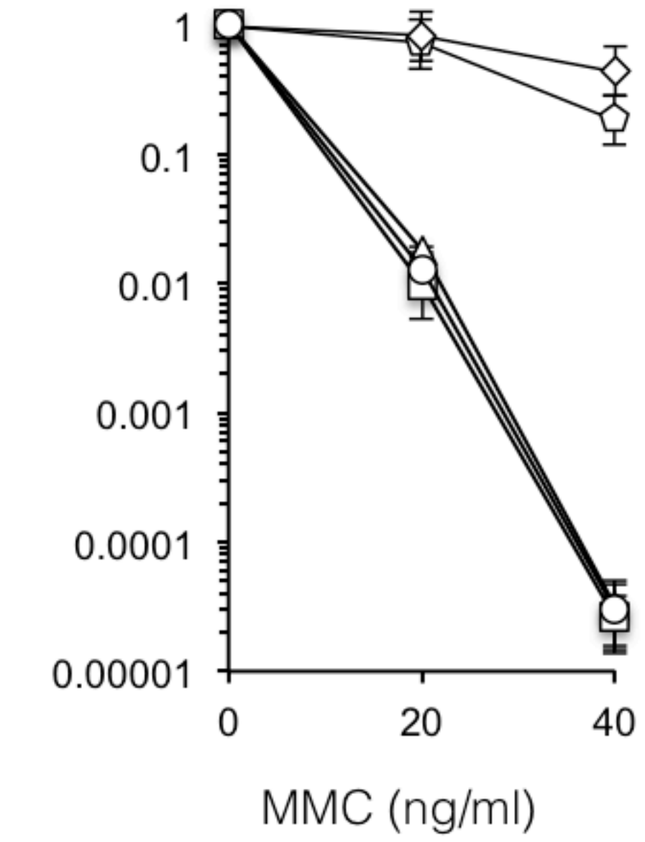
◇ *wild-type*
⬠ *ΔnreA*
△ *ΔuvrA*
□ *ΔuvrA ΔnreA*
○ *ΔuvrA nreA-CΔ10*

E



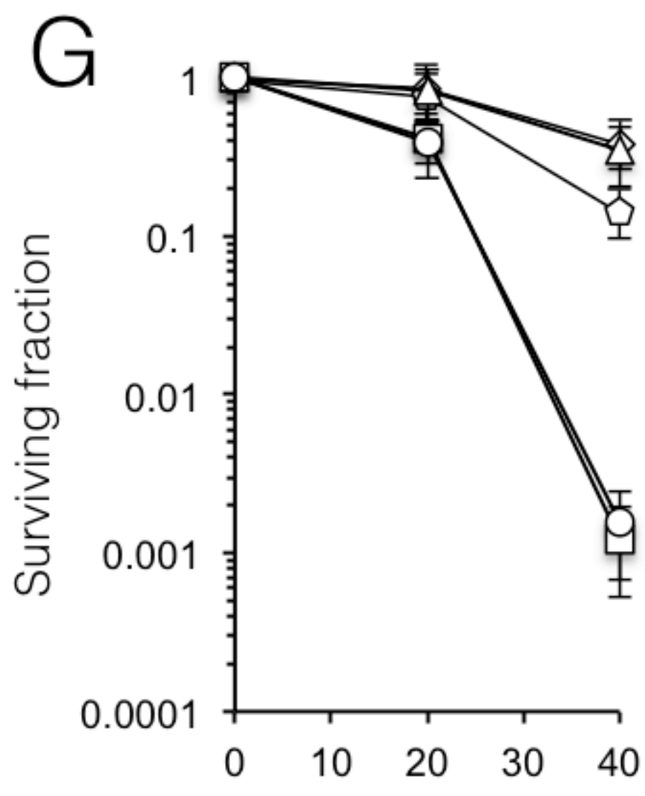
◇ *wild-type*
⬠ *ΔnreA*
△ *ΔuvrB*
□ *ΔuvrB ΔnreA*
○ *ΔuvrB nreA-CΔ10*

F



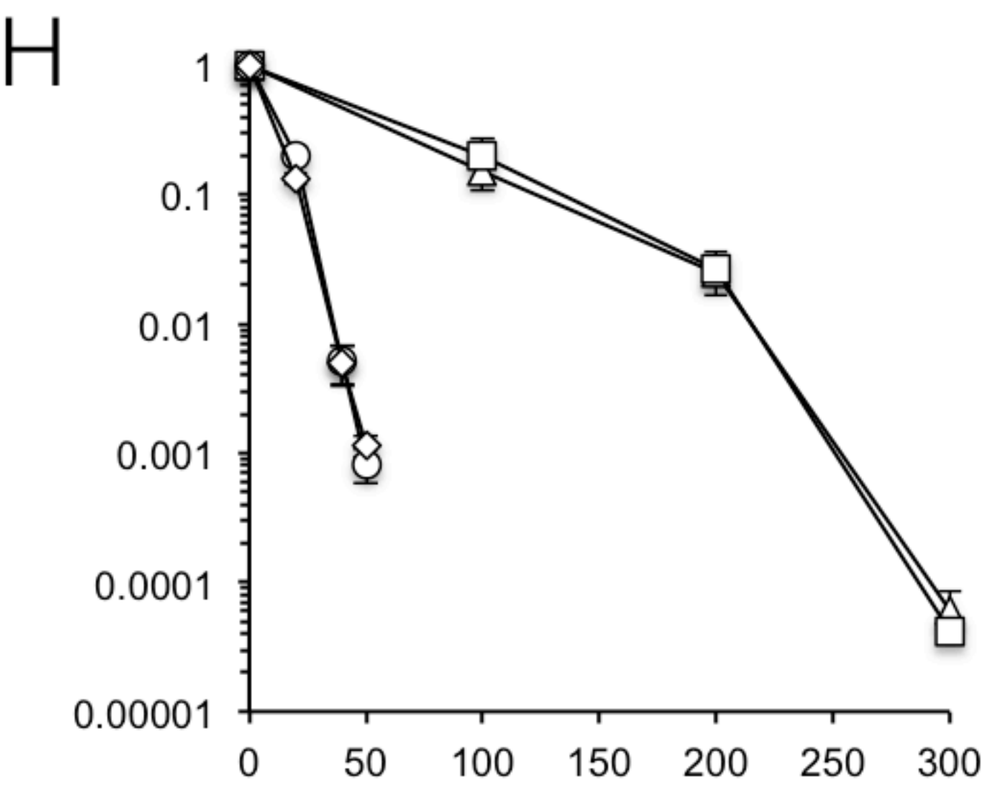
◇ *wild-type*
⬠ *ΔnreA*
△ *ΔuvrC*
□ *ΔuvrC ΔnreA*
○ *ΔuvrC nreA-CΔ10*

G



◇ *wild-type*
⬠ *ΔnreA*
△ *ΔuvrD*
□ *ΔuvrD ΔnreA*
○ *ΔuvrD nreA-CΔ10*

H



△ *wild-type*
□ *ΔnreA*
○ *ΔuvrA*
◇ *ΔuvrA ΔnreA*