G-quadruplexes in *Haloferax volcanii*

Zackie Aktary[1], Anne Cucchiarini[1], Guglielmo Vesco[1,2], Dorian Noury[1], Thomas Jourdain[1], Daniela Verga[3], Pierre Mahou[1], Nicolas Olivier[1], Natalia Valková[4], Otilia Porubiaková[4], Václav Brázda[4], Marie Bouvier[5], Marta Kwapisz[5], Béatrice Clouet-d'Orval[5], Thorsten Allers[6], Roxane Lestini[1], Jean-Louis Mergny[1] & Lionel Guittat[1,7,8]

[1] Laboratoire d'Optique et Biosciences, Ecole Polytechnique, CNRS UMR7645 – INSERM U1182, Institut Polytechnique de Paris, 91120 Palaiseau cedex, France.

[2] Department of Science and High Technology, University of Insubria, 22100 Como, Italy.

[3] Institut Curie, CNRS UMR9187, INSERM U1196, PSL Research University, F-91405 Orsay, France.

[4] Institute of Biophysics of the Czech Academy of Sciences, Královopolská 135, 612 65 Brno, Czech Republic.

[5] Molecular, Cellular and Developmental Biology Unit (MCD), Center for Integrative Biology (CBI; FR3743), University of Toulouse, CNRS, Toulouse, France.

[6] School of Life Sciences, University of Nottingham, Nottingham, United Kingdom.

[7] Université Sorbonne Paris Nord, UFR SMBH, 93000 Bobigny, France.

[8] Corresponding author : email: Lionel.guittat@polytechnique.edu

Keywords: G-quadruplex; archaea; prokaryote; *Haloferax volcanii*; BG4

ABSTRACT

The archaeal domain is a taxonomically rich component of microbial communities that inhabit a wide range of habitats on Earth, including the human body. Phylogenomic analyses have indicated that archaea represent the closest known relatives of eukaryotes, thus suggesting that eukaryotes may have evolved from an archaeal ancestor. G-quadruplex structures (G4), formed by guanine rich sequences, are among the most intensively studied local DNA/RNA structures and regulate key biological processes such as replication and gene expression. A bioinformatics analysis of the genome of the salt-loving archaea *H. volcanii* revealed a large number of potential G4 sequences (PQS). Biophysical analyses showed that a representative panel of these sequences form stable G4 structures under physiological conditions *in vitro*. In addition, immunofluorescence experiments using the G4-specific antibody, BG4, detected G4s *in vivo* at the single-cell level with super-resolution microscopy. Moreover, we directly visualized G4 in exponentially growing or stationary cells both at the DNA and RNA levels. G4s were also observed in the RNA and DNA of the hyperthermophile archaeon *T. barophilus*. Finally, we identified helicases potentially involved in G4 unfolding. Together, with *H. volcanii* as a new model, our work helps to fill the gap between bacteria and eukaryotic organisms for G4 studies and will aid in uncovering the evolutionary history of G4 structures in the tree of life.

INTRODUCTION

Archaea, which constitute one of the three domains of life distinct from both bacteria and eukaryotes, are a diverse and intriguing group of microorganisms that thrive primarily in extreme environments (*1, 2*). Unlike bacteria and eukaryotes, archaea possess unique cell membrane lipids, usually ether-linked isoprenoids, and their cell walls lack peptidoglycan, a key component of bacterial cell walls (*3*). Archaea have contributed precious insights into the evolutionary history of life on Earth. The discovery of the Asgard archaea has supported the more recent concept of a two-domain tree of life, with eukaryotes emerging from the archaea clade (*4-6*). Their singular characteristics make them an endless topic of interest in fields ranging from microbiology to biotechnology and astrobiology. Many archaea are extremophiles, growing in extreme ecosystems that are hostile to the majority of other life forms, such as in environments with elevated temperatures (thermophiles), extreme acidity or alkalinity (acidophiles and alkaliphiles), high salinity (halophiles) and high pressure (piezophiles) (*3*). Under less extreme conditions, archaea play an essential but comparatively less understood role in the human microbiome, notably in the gastrointestinal tract, skin and oral cavity (*7, 8*). Archaea play crucial roles in various biogeochemical cycles, notably the carbon and nitrogen cycles. They are key players in methanogenesis, a critical process in anaerobic environments that contributes to both natural and anthropogenic methane emissions. Some extremophilic archaea are relevant to biotechnology and industrial applications (DNA polymerases, lipases and esterases) (*9*). Although distinct from eukaryotes, they share more molecular and genetic features with eukaryotic organisms than bacteria (e.g. histones and chromatin-

like structures, DNA and RNA polymerases), highlighting their evolutionary significance (*10, 11*). The study of archaea contributes significantly to our understanding of fundamental processes (replication, transcription, translation and repair) and is essential for shedding light on our evolutionary history (*4, 12*).

In addition to the DNA double-helix structure, genomic sequences may adopt non-canonical DNA structures, at least transiently (*13-15*). Among these structures, G-quadruplexes, known as G4s, are unique nucleic acid structures that can be found in DNA and RNA molecules (*16, 17*). They are characterized by the stacking of guanine (G) bases into a four-stranded, square planar arrangement, the G-quartet. In the last decades, G-quadruplexes have gained considerable attention in molecular and genome biology due to their possible roles in various cellular processes and their potential as therapeutic targets (*18-20*). These atypical structures adopt different topologies depending on factors such as the sequence, loop length, and ionic conditions. Monocations, such as potassium (K^+) or sodium (Na^+), can impact quadruplex structure, stability and biological functions (*21-23*). G4s have been implicated in a wide range of biological processes, including transcription (ref?), telomere maintenance (*24*), DNA replication (*25, 26*) and repair (*27, 28*), and RNA processing and translation (*29, 30*). As regulatory elements, G4s act as molecular switches, modulating gene expression or protein-DNA associations (*31, 32*). Recent findings suggest that G4s may participate in epigenetic regulation that contributes to the cell-type specificity of the transcriptome (*33, 34*) and higher-order chromatin structures (*31, 35*). These unusual structures have emerged as potential therapeutic targets for various diseases, including cancer and viral and neurodegenerative disorders (*20, 36, 37*). A very large panel of ligands that stabilize G4s (G4 ligands or

G4L) have been identified and some of them are being investigated for their therapeutic potential (*18, 38-40*).

While extensively investigated in eukaryotes, the presence and functional significance of G4s in archaea have remained largely unexplored (*41*). Furthermore, the evolution of G4s is still largely unexplored. In this article, we have investigated the existence and potential implications of G-quadruplex structures in archaeal genomes, particularly in *Haloferax volcanii*, the salt-loving archaea model (*42*). First, we analyzed differences in the presence and frequency of potential G4-forming sequences (PQS) in the *H. volcanii* genome using the G4Hunter tool. We then experimentally validated G4 formation of sequences identified by bioinformatics using a range of methods. After studying G4 formation *in vitro*, we now demonstrate, for the first time, the presence of G4 in archaeal cells using the G4-specific antibody BG4. These immunofluorescence experiments reveal that G4s are present in both DNA and RNA, and that the presence of these non-B structures can be modulated by a G4-specific ligand or by eukaryotic-like helicases.

RESULTS

PQS in *Haloferax volcanii* genomic DNA

To assess the potential for G4 formation in the *H. volcanii* genome, we first used the G4Hunter algorithm to determine the presence of PQS both on chromosomal DNA (chr) and the four mini-chromosomes (pHV1-4). Using standard values for G4Hunter (*i.e.*, a window size of 25 nucleotides and threshold score of 1.2), we found 5812 PQS in the entire genome of *H. volcanii*. The GC content is above 50% for all *H. volcanii* genomes, with a maximum for the main chromosome and pHV3 (around 66%), followed by pHV4 (61%) then pHV2 (56%) and pHV1 (55%). Total PQS counts, percentage of GC and PQS frequencies characteristics are summarized in Figure 1. The vast majority of these PQS are present in the chromosome (4406), then in the two mini-chromosomes (723 and 605, respectively, for pHV4 and pHV3). The pHV1 contains 68 PQS and the smallest plasmid, pHV2, only 10 PQS. The highest PQS frequencies were found in the main chromosome and pHV2 (PQS around 1.55 per kb) followed by pHV3 and pHV4 (PQS 1.3 and 1.1, respectively). In contrast, the lowest PQS frequency was found in pHV1 (PQS = 0.8). It is worth noting that the plasmid pHV2 has been cured from the laboratory strain (H26). Moreover, in this strain, pHV4 is inserted in the main chromosome, leading to an additional and functional origin of replication (*42, 55*). As for most organisms studied so far, the vast majority (98.5% here) of PQS had a score between 1.2 and 1.4, a relatively low G4Hunter score. This finding applies to all genomes tested to date, including viruses, bacteria, archaea and eukaryotes. There is a correlation in almost all organisms between GC content and the number/frequencies of PQS (*41*). In *H. volcanii*, we investigated whether there was

such a correlation between the GC content of each genomic element (chromosome + mini-chromosomes) and the PQS frequency. We excluded pHV2 from our analysis because its size is very small (6kb) compared to the other components of the genome. We observed a clear correlation between GC content and PQS frequency between the 4 genomic elements (Chr, pHV1, pHV3, pHV4) (Supplementary Material 1).

Experimental evidence for G4 formation *in vitro*

The accuracy of G4Hunter is strong for high scores and slightly inferior for scores between 1.2 and 1.5 (*56*). Since we observed that the majority of G4Hunter scores in *H. volcanii* were between 1.2 and 1.4, it was essential to verify experimentally that some of these identified sequences can adopt the G4 structure *in vitro*. To this end, we selected a panel of 36 sequences, containing four consecutive blocks of at least two guanines, distributed both on the chromosome and on the mini-chromosomes. To demonstrate the formation of the G4, we used a variety of methods (*57*). FRET-MC enables the testing of several sequences in a parallel fashion (*44*). This approach is used to measure the ability of a test sequence to compete with the G4 ligand PhenDC3, which is highly G4-selective, for binding to a fluorescent G4-containing probe, F21T. Binding of PhenDC3 to the F21T increases its T_m. If the test sequence is forming a G4, PhenDC3 will bind to both the test sequence and the F21T, resulting in a less significant change in T_m. Positive and negative controls (sequences known to form or not form G4s) were used for comparison. As shown in Figure 2A, 25 of the 36 sequences behaved as efficient competitors, arguing in favor of G4 formation for the majority of motifs tested.

As several techniques are recommended to definitively demonstrate G4 formation (*57*), we used an alternative approach and investigated whether the fluorescence emission of G4 probes such as thioflavin T (*45*) and NMM (*58*) was increased in the presence of potential G4-forming *H. volcanii* sequences. When compared to the negative and positive controls, most sequences induced an increase in fluorescence emission of both G4 light probes (Figure 2B and 2C). Finally, we performed additional spectroscopic experiments: isothermal difference spectra (IDS) and circular dichroism (CD). The aim of IDS experiments is to compare the absorbance characteristics of the same oligonucleotide in its folded and unfolded forms. The arithmetic difference between the two corresponding spectra results in the isothermal difference spectra. Unfolding can be obtained by excluding stabilizing cations. G4s display a negative peak around 295 nm and a positive peak around 273 nm (Supplementary Figure S1). Most spectra obtained by CD were indicative of parallel G4 formation (Supplementary Figure S2). Altogether, these experiments confirmed that the vast majority of the 36 *H. volcanii* sequences tested are able to fold into G4 conformations (Supplementary Table 4).

G4s are present in archaea *in vivo*

We next aimed to validate our bioinformatic and *in vitro* results by demonstrating that G4s were present in archaea cells *in vivo*. Using 3D-SIM super-resolution imaging, we performed immunofluorescence experiments using the G4-specific single-chain antibody (scFv) BG4 on fixed *H. volcanii* cells (*59*). Given the density of G4 foci, we needed a higher resolution than conventional fluorescence microscopy techniques (250nm) and 3D-SIM microscopy offers a resolution of the order of 100 nm (*60, 61*).

BG4 staining of exponentially growing *H. volcanii* cells revealed the presence of distinct, punctate foci distributed evenly throughout the cells while no staining was detected in the absence of BG4 (Figure 3A and B). These foci were present in *H. volcanii* cells in both exponential and stationary phases of growth, albeit to a significantly higher degree in exponential phase, suggesting the facilitated formation of G4 during genomic DNA replication. Therefore, we use exponential phase cells for our subsequent experiments (Figure 3C and S3).

Using transcriptomic and *in vitro* analyses, Guo and Bartel showed that G4 in RNA (rG4s) are depleted in the *E. coli* genome, most likely due to the lack of an efficient machinery to unwind them (*62*). To determine whether the G4s present in *H. volcanii* were being formed in the RNA, DNA or both, we treated the cells with RNase A before staining with the BG4 antibody. While we were able to detect G4 structures in *H. volcanii* cells after RNase treatment, the G4 signal was decreased approximately 1.7-fold compared to untreated cells (Figure 3D and S4). To confirm that our results were not specific to *H. volcanii*, we performed similar experiments using another archaea: *Thermococcus barophilus*, a hyperthermophilic archaea isolated from a deep-sea hydrothermal vent (*49*). Similar to *H volcanii*, G4s were detected in both exponentially growing and stationary phase *T. barophilus* cells, with a significant increase in signal intensity in exponential phase cells (Supplemental Figure S5). Furthermore, treatment of *T. barophilus* cells with RNase A resulted in a significant decrease in G4 staining compared to untreated controls (Supplemental Figure S6).

Modulation of G4s *in vivo* by a G4-ligand or helicases.

We subsequently treated *H. volcanii* cells with PhenDC3, the G4-ligand already used in our *in vitro* assay. Cells were treated with PhenDC3 for two hours and subsequently processed for immunofluorescence analysis, where we observed a statistically significant increase in G4 staining in PhenDC3 treated vs. untreated cells (Figure 4A and S7).

To further characterize the presence and formation of G4s in our model archaeon, we aimed to identify potential helicases that may be involved in unwinding G4s in *H. volcanii.* As such, we used different *H. volcanii* strains that lacked the expression of specific helicases, focusing on *H. volcanii* helicases whose orthologues had previously been proposed to regulate G4 dynamics *in vitro* and/or *in vivo,* such as Hef and Dna2 and Rad3a/3b. Genetic and imaging studies have revealed the central role of the archaeal helicase/nuclease Hef in repair of arrested replication forks. (*63-65*). In addition to the Hef knock-out strain already available, we generated knock-out strains for Dna2 and Rad3a/3b to study their putative functions regarding G4 unwinding. Immunofluorescence experiments showed that in *H. volcanii* strains lacking these helicases, there was a significant increase in DNA G4 staining compared to the wild-type strain (1.5-, 1.8-, and 1.5-fold for ∆Hef, ∆Dna2 and ∆Rad3a/Rad3b, respectively; Figure 4B and S8).

Discussion

Despite its designation as the third domain of life in the mid 1970s, molecular and cellular biology studies focusing on archaea have been relatively limited, compared to the two other domains, bacteria and eukarya. Yet, archaea are intriguing given that while they are morphologically similar to bacteria (*i.e.*, single-celled organisms), enzymatically, archaea more closely resemble eukaryotes (*3*). While many archaea are found living under harsh conditions, they are also found in almost every environment including the human microbiota. They are one of the most abundant groups of organisms on the planet with important roles both in the carbon and the nitrogen cycle. Interestingly, methanogenesis, which produces the greenhouse methane gas as a by-product, occurs only in Archaea. Moreover, recent analyses support a close relationship between eukaryotes and Asgard archaea. Phylogenomic analyses have established these archaea as the closest prokaryotic relatives of eukaryotes, shedding light on the initial events of eukaryogenesis (*6, 66*). There is therefore a need to understand the biology of these micro-organisms, as they contribute significantly to the life cycle and our evolutionary history. Our work is the first to demonstrate that G4s, atypical nucleic acid structures, are present in archaea *in vivo*. DNA can adopt a wide range of non-canonical structures, which may be biologically important (20, 35, 36). These structures are implicated in human pathologies such as neurological disorders and cancer and have been proposed as promising drug target for the design of therapeutics against various pathogens (18, 37-39). Using publicly available annotated archaeal genomes and using G4-Hunter algorithm, we were able to predict that G4s could be formed in the genome of every archaeal species examined (*67*). In this study, we focused primarily on *H. volcanii*, a

halophilic mesophile archaeon, since: 1) it is relatively simple to cultivate and manipulate experimentally (an aerobic organism growing in 2.5M NaCl medium); 2) its genome has a high GC content (66%); and 3) according to our G4-Hunter algorithm, its genome has approximately 1.55 putative quadruplex forming sequences (PQS) per kb (in comparison to the human genome, which as a PQS score of 1.2). Our *in vitro* experiments using various sequences from the *H. volcanii* genome with a G4-Hunter score of at least 1.2 demonstrated that the vast majority of tested sequences were prone to forming G4 *in vitro.* Importantly, these tested sequences were derived from not only from the *H. volcanii* main circular chromosome, but they were also taken from the two mega plasmids found within the strain used in our study (pHV1 and pHV3). This result suggests that G4s are found in all components of the *H. volcanii* genome, and are not limited to one specific element.

To unequivocally show that G4s were present in archaea, we performed immunofluorescence experiments on *H. volcanii* cells in exponential and stationary phase. These analyses clearly demonstrated an enrichment in G4s in exponential phase cells. This difference in G4 content can most likely be attributed to the fact that exponentially dividing cells are actively replicating their DNA content as they undergo further rounds of cell growth. As such, there would be a higher number of G4s in these cells compared to stationary phase cells, which have reached a plateau and divide much less frequently. We confirmed these results using the hyperthermophilic archaeon, *T. barophilus*, which grows under anaerobic conditions and metabolizes sulfur (*49*). Importantly, these results indicate that G4s can be formed in organisms that grow in considerably different conditions, implying an underlying prevalence of

these structures in not only archaea, but presumably throughout the three domains of life.

G4s were present in both the DNA and RNA in both *H. volcanii* and *T. barophilus*, as indicated by a decreased G4 signal in cells following treatment with RNase A. This result is consistent with observations in eukaryotes, where G4s are known to also be formed not only in DNA, but also in RNA (*29, 68*). In higher organisms, DNA G4s are implicated in regulating a number of fundamental processes, such as DNA replication, transcription and telomere maintenance (*16*), whereas RNA G4s are thought to play a role in regulating translation and alternative splicing (*29, 30*). Interestingly, G4s are thought to regulate these processes both positively, through the recruitment of various initiating factors, or negatively, by acting as physical roadblocks, preventing access to the appropriate molecular machinery (*69*). The presence of G4s in both the RNA and DNA in archaea would suggest that these structures play similar functions in a diverse array of evolutionarily distinct organisms, underlying the fundamental nature and importance of G4s.

Having confirmed the presence of G4s in archaea, we aimed to identify potential players that are implicated in G4 stability/unfolding, using various helicase mutant strains of *H. volcanii*. Hef is a unique member of the XPF/MUS81 family of helicases/translocases and is an orthologue of the human Fanconi anemia protein FANCM, which has previously been implicated in destabilization of G4s, as human carcinoma cell lines lacking expression of FANCM were more sensitive to pyridostatin (PDS), a G4 ligand (*63, 64, 70*). Dna2 is a helicase/endonuclease that has been shown to bind to and unwind G4s, primarily in telomeres both in yeast and human cell lines (*71, 72*). Rad3a and Rad3b are putative orthologs of human XPB and XPD proteins,

respectively, which are components of the TFIIH general transcription factor (*73*). These proteins have both been shown to bind G4s and XPB acts as a G4 DNA helicase (*74*). Interestingly, our results have shown that loss of either of these helicases resulted in increased DNA G4s. This raises the intriguing possibility that there are multiple helicases and enzymatic pathways that are implicated in regulating G4 stability and dynamics in archaea. Furthermore, it remains unclear whether these helicases cooperate in unwinding G4s, which may be possible, as the most notable increase in G4 intensity was approximately 1.8-fold in ΔDna2 cells. From a cell biological perspective, having multiple helicases being implicated in regulating G4 unfolding would be advantageous, especially when one particular helicase is absent. These molecular machineries would work together to prevent an excess of G4s, which could be deleterious to cell health, growth and survival.

The fact that orthologous human and archaeal helicases are involved in regulating G4 unfolding, and that similar to human cells, G4s can be found in the RNA and DNA in archaea, would suggest a level of evolutionary conservation with respect to the formation and regulation of G4s. This finding points to the critical importance and fundamental role of these secondary structures in regulating basic fundamental processes in cell and molecular biology. Furthermore, this conservation has exciting implications in the field of evolutionary biology, and suggests that the molecular tools that have evolved over time to manage and regulate G4 formation in eukaryotes and specifically higher eukaryotes existed in archaea. The question of when these molecular machineries evolved during evolution is more intriguing considering that previous studies suggested that in *E. coli*, RNA G4s are depleted, primarily due to the lack of the appropriate molecular machinery to unfold these structures. Indeed,

exogenous expression of RNA G4 forming sequences in *E. coli* resulted in decreased cell growth, suggesting that the cells did not have the capacity to properly resolve these structures (*62*). It should be noted that in this work, the presence of G4s *in vivo*, by microscopy, was not assessed.

It is now widely accepted that G4s can play antagonistic roles: helping some molecular processes, but also interfering with them. The origin of these dual roles can be investigated by studying prokaryotic organisms. While the formation of G4s in bacteria has been documented, the study of their functions is still very limited compared to eukaryotes (*75*). Until now, there have been virtually no data on the detection and function of G4s in archaea. This work opens up new perspectives for understanding the relevance of G4s and how they influence molecular processes such as replication, transcription, translation and repair. The identification of potential helicases involved in G4 unwinding provides a tool for studying their impact, although the significance of these proteins *in vivo* remains to be better defined.

Overall, our work is the first, unequivocal demonstration of G4 formation in archaea, an exciting result that opens the door to further studies that will help to answer fundamental questions about the evolution of organisms and the molecular mechanisms regulating DNA structure, dynamics and the regulation of gene expression.

MATERIAL AND METHODS

Analytical process

All sequences were analyzed using the G4Hunter Web tool (http://bioinformatics.ibp.cz/#/analyse/quadruplex), which can read National Center for Biotechnology Information (NCBI) IDs. Unless specified otherwise, parameters for G4Hunter were set to 25 nucleotides for window size and a threshold score of at least 1.2. This threshold appears as a reasonable compromise, giving few false positives (sequences not forming a G4 despite a G4Hunter score above threshold) and false negatives (sequences able to form a stable G4 despite having a G4Hunter score below threshold). Scores above 1.2 correspond to sequences having a relatively high guanine content and likely to form stable G4s. To rank sequences based on score, motifs were binned in five intervals covering the G4Hunter scores 1.2–1.4, 1.4–1.6, 1.6–1.8, 1.8–2.0 and >2.0 (*43*). The higher the score, the higher the probability to form a stable G4.

Oligonucleotide sequences

Oligonucleotides were purchased from Eurogentec (Belgium) and used without further purification. F21T, the fluorescently labeled reference quadruplex, negative and positive controls, as well as the tested sequences were prepared at 100 µM strand concentration in ddH_2O. Sequence details are provided in Supplementary Tables 1, 2 and 3. All oligonucleotides were annealed (95°C for 5 min and slowly cooled to room temperature) in the corresponding buffer before measurements.

Circular dichroism

3 µM oligonucleotide solutions were annealed in K100 buffer (10 mM lithium cacodylate, 100 mM KCl, pH 7.2). CD spectra were recorded on a J-1500 spectropolarimeter (JASCO, France) at room temperature (25°C), using a scan range of 350–210 nm, a scan rate of 100 nm/min and averaging 3 accumulations.

FRET-melting competition assay

FRET melting competition (FRET-MC) experiments were performed in 96-well plates using a CFX96 qPCR Real-Time PCR (Biorad). The tested sequences (or competitors) were annealed at 7.5 μM in K10 buffer (10 mM KCl, 10 mM lithium cacodylate, 90 mM LiCl, pH 7.2) and F21T was annealed at 5μM in K10. Each well contained 3 μM competitors, 0.2 µM fluorescent oligonucleotide F21T with or without 0.4 µM G4 ligand (PhenDC3) in K10 buffer, for a total volume of 25 µL. Samples were kept at 25°C for 5 min, then the temperature was increased by 0.5°C per minute until 95°C, the FAM channel was used to collect the fluorescence signal. The melting temperature (T_m) of an oligonucleotide is defined as the temperature where 50% of the oligonucleotide is unfolded and corresponds to the normalized value of 0.5. ΔT_m is determined by substracting the T_m containing PhenDC3 to the T_m in absence of PhenDC3. The S-Factor, a measure for the strength of G4 formation, has been previously described (*44*). The FRET-MC assay uses a 96-well plate as a sample holder, which allows to process 48 sequences simultaneously. Each experimental condition was tested in duplicate on two separate plates.

Isothermal difference spectra (IDS)

Absorbance spectra were recorded on a Cary 300 spectrophotometer (Agilent Technologies, France) (scan range: 330-220 nm; scan rate: 600 nm/min; automatic baseline correction). 3 µM oligonucleotide solutions were annealed in Li-CaCo10 buffer (10 mM lithium cacodylate, pH 7.2). Absorbance spectra were first recorded at 25°C in the absence of any stabilizing cation. 1M KCl was added after recording the first spectrum, to a final potassium concentration of 100 mM. The second UV-absorbance spectrum was recorded after 15 min of equilibration. IDS corresponds to the arithmetic difference between the initial (unfolded) and second (folded, thanks to the addition of K^+) spectra, after correction for dilution.

G-Quadruplex fluorescent light-up probes

Thioflavin T (ThT) was used as previously described (*45*). The tested oligonucleotides were annealed at 7.5 μM in K100 buffer. The oligonucleotides and the ThT (dissolved in milli Q water) were added in a microplate at 3 and 2 μM respectively, to reach a final volume of 100μL in K100 buffer. The plate was shaken for 10 sec and was incubated for 10 min at room temperature. Fluorescence intensity was collected at 490 nm after excitation at 420 nm in a TECAN M1000 pro plate reader. The fluorescence signal was normalized to the oligonucleotide length. Each condition was done in duplicate on different plates.

NMM (N-methyl mesoporphyrin IX) was used under the same condition as ThT, except that fluorescence intensity was collected at 610 nm after excitation at 380 nm in a TECAN M1000 pro plate reader.

Strains, plasmids and growth conditions

H. volcanii cultures using enriched Hv-YPC media were grown at 45°C, as described previously (*46*). Hv-YPC contained (per liter) 144 g of NaCl, 21 g of $MgSO_4$.7H2O, 18 g of $MgCl_2$.6H2O, 4.2 g of KCl, and 12 mM Tris–HCl (pH 7.5), Yeast extract (0.5%, wt/vol; Difco), 0.1% (wt/vol) peptone (Oxoid), and 0.1% (wt/vol) Casamino Acids (Difco). For experiments using the G4L PhenDC3, overnight cultures were diluted to OD600 of 0.1 and incubated at 45°C. PhenDC3 (Sigma) was added (to a final concentration of 10 μM) after one hour of growth, and the treatment was continued for two hours at 45°C.

Cells of *Thermococcus barophilus* (TbΔ517; UBOCC-M-3300; (*47*)) were grown at 85°C under atmospheric pressure in anaerobic conditions (50 mL penicillin vials) in Thermococcales rich medium (TRM; Zeng et al., 2009 (*48*)) supplemented with elemental sulphur (0.25 g/L; Marteinsson (*49*)). To determine growth, cells were counted using a Thoma counting chamber (0.01 mm depth, Weber, England). For exponential phase, cells were grown for 6h after inoculation from saturated pre-culture (1×107 cells/mL) and for stationary phase for 18h (1×108 cells/mL).

Construction of strains with deletions of Rad3 and Dna2 helicases

To generate the Δdna2 construct (strain H431), the upstream regions (US) of dna2 and the down-stream regions (DS) of dna2 were cloned into pTA131 to generate the pTA417 plasmids, which was used to construct the deletion strain by a gene knockout system (44). The dna2 deletion was generated in the H53 strain (ΔpyrE2 ΔtrpA) (*46*). The US and DS regions were generated by PCR on H26 genomic DNA (ΔpyrE2) using specific primers. The template DNA used was isolated from genomic

DNA libraries of H. volcanii using internal primers with BamHI sites and external primers with XbaI and KpnI sites, inserted at KpnI and XbaI sites in pTA131 plasmid.

To generate the Δrad3a and Δrad3b constructs (strains HvLG1 and HvLG2, respectively), the upstream regions (US) of *rad3a/b* and the down-stream regions (DS) of *rad3a/b* were cloned into pTA131 to generate the pLG1/2 plasmids, which were used to construct the deletion strains by a gene knockout system (*50*). The rad3a/b deletion was generated in the H26 strain (ΔpyrE2) (*46*). The US and DS regions were generated by PCR on H26 genomic DNA (ΔpyrE2) using specific primers. Each PCR product contained 30 bp homology with adjacent fragments for sequence and ligation independent cloning (SLIC) (*51*). Following the SLIC method, PCR fragments and linearized plasmid were digested by T4 DNA polymerase exonuclease activity for 45 min at 22°C to generate 3'-single-stranded extremities, then all amplification products were mixed in a 1:1:1 molecular ratio and incubated 30 min at 37°C prior to transformation into E. coli XL1-blue cells. The presence of the correct inserts was tested by PCR amplification on white colonies using specific primers. The sequence of selected plasmids, pLG1 and 2, were further confirmed by sequencing on both strands. Both plasmids were used to transform strain H26 (ΔpyrE2) using the pop-in/pop-out method as described previously (*50*). The absence of the *rad3a* and *rad3b* genes was tested by colony lift on 100 colonies from the pop-out plates using a probe targeting *rad3a* or *rad3b*. The digoxigenin (DIG)-labeled probes were generated by PCR on H26 genomic DNA using primers and the PCR DIG labeling Mix from Roche. Probe hybridization was detected using the DIG Luminescent Detection kit (Roche) and a ChemiDoc MP (BioRad). All the oligonucleotides used for the different constructions

are listed in Supplemental Table 3. The double Δrad3a Δrad3b mutant (strain HvLG3) was constructed by deleting rad3b in the Δrad3a mutant as previously described.

In vivo visualization of G4 structures: 3D-Structured Illumination microscopy (3D-SIM)

H volcanii cells were fixed with 2 % formaldehyde (FA) for 30 min at 45°C followed by the addition of glycine (0.5M) for 20 minutes under agitation. Cells were washed with PBS, permeabilized with 60% ethanol, washed and resuspended in 1 ml PBS for cultures with an OD600 of 0.2. 20 µl of cell solution was spotted on poly-L-Lysine coated coverslips for 30 minutes, washed with PBS and incubated with 100 µg/ml RNaseA (ThermoFisher) for 30 min at 37°C. Coverslips were then washed with PBS and blocked with block solution (PBS +b5% goat serum + 0.1% Tween) for 1 hour at room temperature. Next, coverslips were incubated with the BG4 antibody (10 nM) for 1 h . Coverslips were then washed with PBS and incubated with a mouse anti-FLAG antibody (1/1000) for 1 h. After washing with PBS, slides were incubated with anti-Mouse IgG-DyLight (1/1000) for 40 min. All incubations were done at room temperature and all antibodies were prepared in antibody solution (PBS + 1% goat serum + 0.1% Tween). Slides were washed two times with PBS-0.1% Tween, then incubated with Yoyo1 (Invitrogen) for 10 minutes at room temperature and washed three times with PBS. Images were acquired on a DeltaVision OMX SR imaging system from GE Healthcare (Velizy, France) with an Olympus PlanApo N 60× 1.42 NA oil immersion objective. The quality of the raw SI data was checked by the SIMcheck suite of plugins available for Fiji software prior to SI reconstruction performed with the softWoRx software from Applied Precision. Image analysis was performed using

custom scripts in Napari (*52*). Briefly, individual cell outlines were segmented using the nucleic acid signal and Cellpose's pretrained models (*53*). In each cell, the integrated signal and other features were determined using scikit-image package (*54*). The whole analysis pipeline will be made available with the final version of this article.

BG4 purification

The BG4-encoding plasmid (Plasmid #55756, Addgene), was transformed into *Escherichia coli* strain Rosetta 2 pLys (Novagen), followed by culturing in 2xYT and kanamycin (50 µg/ml) and chloramphenicol (34 µg/ml) at 37°C and 250 RPM. When OD600 reached 0.6–0.8 the temperature was lowered to 20°C, and induction was initiated with 0.5 mM IPTG. After 15 h of induction at 20°C, the bacterial culture was harvested by spinning at 4,000 g for 30 min. The resulting pellet was resuspended in 8 ml TES buffer 1 (50 mM Tris-HCl pH 8.0, 1 mM EDTA pH 8.0, 20% sucrose) per 100 ml volume expression culture. The mixture was put on ice for 10 min. Next, 12 ml of TES buffer 2 (10 mM Tris-HCl pH 8.0, 0.2 mM EDTA pH 8.0, 4% sucrose, with benzonase nuclease and 2 mM MgSO4) per 100 ml volume expression culture was added and the mixture was put on ice for 15 min. Cells were centrifuged at 4°C for 20 min at 4000 g. The supernatant was collected and filtered (0.22 µm). The sample was deposited on a washed and PBS-preequilibrated HisTrap HP column (Cytiva). The resin was then washed with 20 column volumes (CV) of high salt wash buffer (PBS, 350 mM NaCl, 10 mM imidazole pH 8.0) and 20 CV low salt wash buffer (PBS, 10 mM imidazole pH 8.0). The BG4 protein was eluted with elution buffer (PBS, 500 mM imidazole pH 8.0), buffer-exchanged to PBS (Econo-Pac 10DG Column, Biorad) and concentrated using 50 ml centrifugal devices with 10 kDa cut-off (Millipore). Protein

quality was assessed by running SDS-PAGE and visualized with Coomassie staining. Final protein concentration was calculated based on absorbance measurements at 280 nm.

References

1. C. R. Woese, O. Kandler, M. L. Wheelis, Towards a natural system of organisms: proposal for the domains Archaea, Bacteria, and Eucarya. *Proc Natl Acad Sci U S A* **87**, 4576-4579 (1990).
2. B. J. Baker *et al.*, Diversity, ecology and evolution of Archaea. *Nat Microbiol* **5**, 887-900 (2020).
3. M. van Wolferen, A. A. Pulschen, B. Baum, S. Gribaldo, S. V. Albers, The cell biology of archaea. *Nat Microbiol* **7**, 1744-1755 (2022).
4. S. Gribaldo, C. Brochier-Armanet, Evolutionary relationships between archaea and eukaryotes. *Nat Ecol Evol* **4**, 20-21 (2020).
5. G. Tahon, P. Geesink, T. J. G. Ettema, Expanding Archaeal Diversity and Phylogeny: Past, Present, and Future. *Annu Rev Microbiol* **75**, 359-381 (2021).
6. A. Spang *et al.*, Asgard archaea are the closest prokaryotic relatives of eukaryotes. *PLoS Genet* **14**, e1007080 (2018).
7. C. M. Chibani *et al.*, A catalogue of 1,167 genomes from the human gut archaeome. *Nat Microbiol* **7**, 48-61 (2022).
8. P. Geesink, T. J. G. Ettema, The human archaeome in focus. *Nat Microbiol* **7**, 10-11 (2022).
9. D. Aparici-Carratalá, J. Esclapez, V. Bautista, M. J. Bonete, M. Camacho, Archaea: current and potential biotechnological applications. *Res Microbiol* **174**, 104080 (2023).
10. K. M. Stevens *et al.*, Histone variants in archaea and the evolution of combinatorial chromatin complexity. *Proc Natl Acad Sci U S A* **117**, 33384-33395 (2020).
11. K. M. Stevens, T. Warnecke, Histone variants in archaea - An undiscovered country. *Semin Cell Dev Biol* **135**, 50-58 (2023).
12. D. Kazlauskas, M. Krupovic, J. Guglielmini, P. Forterre, Č. Venclovas, Diversity and evolution of B-family DNA polymerases. *Nucleic Acids Res* **48**, 10142-10156 (2020).
13. G. Wang, K. M. Vasquez, Dynamic alternative DNA structures in biology and disease. *Nat Rev Genet* **24**, 211–234 (2023).
14. K. D. Makova, M. H. Weissensteiner, Noncanonical DNA structures are drivers of genome evolution. *Trends Genet* **39**, 109-124 (2023).
15. W. M. Guiblet *et al.*, Non-B DNA: a major contributor to small- and large-scale variation in nucleotide substitution frequencies across the genome. *Nucleic Acids Res* **49**, 1497-1516 (2021).
16. D. Varshney, J. Spiegel, K. Zyner, D. Tannahill, S. Balasubramanian, The regulation and functions of DNA and RNA G-quadruplexes. *Nat Rev Mol Cell Biol* **21**, 459-474 (2020).
17. J. Spiegel, S. Adhikari, S. Balasubramanian, The Structure and Function of DNA G-Quadruplexes. *Trends Chem* **2**, 123-136 (2020).
18. A. Abiri *et al.*, Unlocking G-Quadruplexes as Antiviral Targets. *Pharmacol Rev* **73**, 897-923 (2021).
19. S. Debbarma, P. C. Acharya, Targeting G-Quadruplex DNA for Cancer Chemotherapy. *Curr Drug Discov Technol* **19**, e140222201110 (2022).
20. N. Kosiol, S. Juranek, P. Brossart, A. Heine, K. Paeschke, G-quadruplexes: a promising target for cancer therapy. *Mol Cancer* **20**, 40 (2021).
21. G. N. Parkinson, M. P. Lee, S. Neidle, Crystal structure of parallel quadruplexes from human telomeric DNA. *Nature* **417**, 876-880 (2002).
22. S. Burge, G. N. Parkinson, P. Hazel, A. K. Todd, S. Neidle, Quadruplex DNA: sequence, topology and structure. *Nucleic Acids Res* **34**, 5402-5415 (2006).
23. Y. Luo *et al.*, A sodium/potassium switch for G4-prone G/C-rich sequences. *Nucleic Acids Res* **52**, 448-461 (2024).
24. T. M. Bryan, G-Quadruplexes at Telomeres: Friend or Foe? *Molecules* **25**, (2020).
25. J. Poulet-Benedetti *et al.*, Dimeric G-quadruplex motifs-induced NFRs determine strong replication origins in vertebrates. *Nat Commun* **14**, 4843 (2023).

26. P. Prorok *et al.*, Involvement of G-quadruplex regions in mammalian replication origin activity. *Nat Commun* **10**, 3274 (2019).
27. K. Sato, P. Knipscheer, G-quadruplex resolution: From molecular mechanisms to physiological relevance. *DNA Repair (Amst)* **130**, 103552 (2023).
28. R. C. Duardo, F. Guerra, S. Pepe, G. Capranico, Non-B DNA structures as a booster of genome instability. *Biochimie* **214**, 176-192 (2023).
29. K. Lyu, E. Y. Chow, X. Mou, T. F. Chan, C. K. Kwok, RNA G-quadruplexes (rG4s): genomics and biological functions. *Nucleic Acids Res* **49**, 5426-5450 (2021).
30. L. Dumas, P. Herviou, E. Dassi, A. Cammas, S. Millevoi, G-Quadruplexes in RNA Biology: Recent Advances and Future Directions. *Trends Biochem Sci* **46**, 270-283 (2021).
31. J. Shen *et al.*, Promoter G-quadruplex folding precedes transcription and is controlled by chromatin. *Genome Biol* **22**, 143 (2021).
32. J. Spiegel *et al.*, G-quadruplexes are transcription factor binding hubs in human chromatin. *Genome Biol* **22**, 117 (2021).
33. S. Lago *et al.*, Promoter G-quadruplexes and transcription factors cooperate to shape the cell type-specific transcriptome. *Nat Commun* **12**, 3885 (2021).
34. K. G. Zyner *et al.*, G-quadruplex DNA structures in human stem cells and differentiation. *Nat Commun* **13**, 142 (2022).
35. Y. Hou *et al.*, Integrative characterization of G-Quadruplexes in the three-dimensional chromatin structure. *Epigenetics* **14**, 894-911 (2019).
36. E. Ruggiero, I. Zanin, M. Terreri, S. N. Richter, G-Quadruplex Targeting in the Fight against Viruses: An Update. *Int J Mol Sci* **22**, (2021).
37. E. Wang, R. Thombre, Y. Shah, R. Latanich, J. Wang, G-Quadruplexes as pathogenic drivers in neurodegenerative disorders. *Nucleic Acids Res* **49**, 4816-4830 (2021).
38. J. Lopes-Nunes, P. A. Oliveira, C. Cruz, G-Quadruplex-Based Drug Delivery Systems for Cancer Therapy. *Pharmaceuticals (Basel)* **14**, (2021).
39. E. Mendes, I. M. Aljnadi, B. Bahls, B. L. Victor, A. Paulo, Major Achievements in the Design of Quadruplex-Interactive Small Molecules. *Pharmaceuticals (Basel)* **15**, (2022).
40. G. Miglietta, J. Marinello, M. Russo, G. Capranico, Ligands stimulating antitumour immunity as the next G-quadruplex challenge. *Mol Cancer* **21**, 180 (2022).
41. G. Marsico *et al.*, Whole genome experimental maps of DNA G-quadruplexes in multiple species. *Nucleic Acids Res* **47**, 3862-3874 (2019).
42. P. Pérez-Arnaiz, A. Dattani, V. Smith, T. Allers, -a model archaeon for studying DNA replication and repair. *Open Biol* **10**, 200293 (2020).
43. V. Brázda *et al.*, G4Hunter web application: a web server for G-quadruplex prediction. *Bioinformatics* **35**, 3493-3495 (2019).
44. Y. Luo, A. Granzhan, D. Verga, J. L. Mergny, FRET-MC: A fluorescence melting competition assay for studying G4 structures in vitro. *Biopolymers* **112**, e23415 (2021).
45. A. Renaud de la Faverie, A. Guédin, A. Bedrat, L. A. Yatsunyk, J. L. Mergny, Thioflavin T as a fluorescence light-up probe for G4 formation. *Nucleic Acids Res* **42**, e65 (2014).
46. T. Allers, H. P. Ngo, M. Mevarech, R. G. Lloyd, Development of additional selectable markers for the halophilic archaeon Haloferax volcanii based on the leuB and trpA genes. *Appl Environ Microbiol* **70**, 943-953 (2004).
47. T. Birien *et al.*, Development of an Effective 6-Methylpurine Counterselection Marker for Genetic Manipulation in Thermococcus barophilus. *Genes (Basel)* **9**, (2018).
48. X. Zeng *et al.*, Pyrococcus CH1, an obligate piezophilic hyperthermophile: extending the upper pressure-temperature limits for life. *ISME J* **3**, 873-876 (2009).
49. V. T. Marteinsson *et al.*, Thermococcus barophilus sp. nov., a new barophilic and hyperthermophilic archaeon isolated under high hydrostatic pressure from a deep-sea hydrothermal vent. *Int J Syst Bacteriol* **49 Pt 2**, 351-359 (1999).
50. G. Bitan-Banin, R. Ortenberg, M. Mevarech, Development of a gene knockout system for the halophilic archaeon Haloferax volcanii by use of the pyrE gene. *J Bacteriol* **185**, 772-778 (2003).

51. M. Z. Li, S. J. Elledge, Harnessing homologous recombination in vitro to generate recombinant DNA via SLIC. *Nat Methods* **4**, 251-256 (2007).
52. N. Sofroniew *et al.* (Zenodo, 2022).
53. C. Stringer, T. Wang, M. Michaelos, M. Pachitariu, Cellpose: a generalist algorithm for cellular segmentation. *Nat Methods* **18**, 100-106 (2021).
54. S. van der Walt *et al.*, scikit-image: image processing in Python. *PeerJ* **2**, e453 (2014).
55. M. Hawkins, S. Malla, M. J. Blythe, C. A. Nieduszynski, T. Allers, Accelerated growth in the absence of DNA replication origins. *Nature* **503**, 544-547 (2013).
56. A. Bedrat, L. Lacroix, J. L. Mergny, Re-evaluation of G-quadruplex propensity with G4Hunter. *Nucleic Acids Res* **44**, 1746-1759 (2016).
57. Y. Luo *et al.*, Guidelines for G-quadruplexes: I. In vitro characterization. *Biochimie*, (2022).
58. N. C. Sabharwal *et al.*, N-methylmesoporphyrin IX fluorescence as a reporter of strand orientation in guanine quadruplexes. *FEBS J* **281**, 1726-1737 (2014).
59. R. Hänsel-Hertsch *et al.*, G-quadruplex structures mark human regulatory chromatin. *Nat Genet* **48**, 1267-1272 (2016).
60. C. Bond, A. N. Santiago-Ruiz, Q. Tang, M. Lakadamyali, Technological advances in super-resolution microscopy to study cellular processes. *Mol Cell* **82**, 315-332 (2022).
61. D. Alsafadi, S. Alsalman, F. Paradisi, Extreme halophilic alcohol dehydrogenase mediated highly efficient syntheses of enantiopure aromatic alcohols. *Org Biomol Chem* **15**, 9169-9175 (2017).
62. J. U. Guo, D. P. Bartel, RNA G-quadruplexes are globally unfolded in eukaryotic cells and depleted in bacteria. *Science* **353**, (2016).
63. R. Lestini *et al.*, Intracellular dynamics of archaeal FANCM homologue Hef in response to halted DNA replication. *Nucleic Acids Res* **41**, 10358-10370 (2013).
64. R. Lestini, Z. Duan, T. Allers, 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 (2010).
65. K. Komori *et al.*, Cooperation of the N-terminal Helicase and C-terminal endonuclease activities of Archaeal Hef protein in processing stalled replication forks. *J Biol Chem* **279**, 53175-53185 (2004).
66. L. Eme *et al.*, Inference and reconstruction of the heimdallarchaeial ancestry of eukaryotes. *Nature* **618**, 992-999 (2023).
67. V. Brázda *et al.*, G-Quadruplexes in the Archaea Domain. *Biomolecules* **10**, (2020).
68. C. K. Kwok, G. Marsico, S. Balasubramanian, Detecting RNA G-Quadruplexes (rG4s) in the Transcriptome. *Cold Spring Harb Perspect Biol* **10**, (2018).
69. J. Robinson, F. Raguseo, S. P. Nuccio, D. Liano, M. Di Antonio, DNA G-quadruplex structures: more than simple roadblocks to transcription? *Nucleic Acids Res* **49**, 8419-8431 (2021).
70. S. Li *et al.*, ERCC1/XPF Is Important for Repair of DNA Double-Strand Breaks Containing Secondary Structures. *iScience* **16**, 63-78 (2019).
71. L. Zheng, Y. Meng, J. L. Campbell, B. Shen, Multiple roles of DNA2 nuclease/helicase in DNA metabolism, genome stability and human diseases. *Nucleic Acids Res* **48**, 16-35 (2020).
72. W. Lin *et al.*, Mammalian DNA2 helicase/nuclease cleaves G-quadruplex DNA and is required for telomere integrity. *EMBO J* **32**, 1425-1439 (2013).
73. J. Kuper, C. Kisker, Three targets in one complex: A molecular perspective of TFIIH in cancer therapy. *DNA Repair (Amst)* **105**, 103143 (2021).
74. L. T. Gray, A. C. Vallur, J. Eddy, N. Maizels, G quadruplexes are genomewide targets of transcriptional helicases XPB and XPD. *Nat Chem Biol* **10**, 313-318 (2014).
75. R. R. Cueny, S. D. McMillan, J. L. Keck, G-quadruplexes in bacteria: insights into the regulatory roles and interacting proteins of non-canonical nucleic acid structures. *Crit Rev Biochem Mol Biol* **57**, 539-561 (2022).

Figure Legends

Fig. 1. GC content, total number of PQS and their frequencies in H. volcanii genome. (A) Schematic presentation of a G-quartet and a G-quadruplex. (B) Total PQS counts, percentage of GC and PQS frequencies characteristics for the main chromosome and mini-chromosomes.

Fig. 2. Evidence for G4 formation with 36 *H. volcanii* sequences. (A) FRET MC results. The S-factor gives is a score concerning G4 formation. The lower the S-factor, the higher the G4 formation. (B) Thioflavin T fluorescence emission and (C) NMM fluorescence emission. Both Thioflavin T and NMM are light-up probes for which fluorescence intensity increases in the presence of G4-forming sequences.

Fig. 3. Immunofluorescence detection of G4s in *H. volcanii in vivo*. (A) Immunofluorescence labelling of G4 in *H. volcanii* using the BG4 antibody. YOYO-1 was used to visualize nucleic acids. (B) Magnification of BG4-stained *H. volcanii* cells. Bar, 5 μm. (C) Immunofluorescence labelling and quantification of total G4 fluorescence intensity in *H. volcanii* exponential and stationary phase cells and in exponential phase cells treated with RNase A (D). Statistical analyses were performed using a Mann-Whitney test. **** $p > 0.0001$. Bar, 5 μm.

Fig. 4. Chemical and genetic modulation of G4s in *H. volcanii*. (A) Quantification of total G4 fluorescence intensity in exponential phase *H. volcanii* cells treated with the G4 ligand, PhenDC3 and in different helicase mutant *H. volcanii* strains (B). Statistical analyses were performed using a Mann-Whitney test. **** $p > 0.0001$. Bar, 5 μm.

Acknowledgements

We thank M. Imezar, H. Becker, L. Mellottee, T Garnier for technical assistance (LOB); JC. Andrau (IGMM) and L. Lacroix (ENS) for helpful discussions.

Funding :

ANR G4Access [ANR-20-CE12-0023]

ANR Morphoscope2 [ANR-11-EQPX-0029]

ANR France BioImaging [ANR-10-INBS-04]

Author contributions:

Conceptualization: LG, JLM, RL

Methodology: NO, PM, AC, DN, TA, VB, TJ, ZA, LG

Investigation: AC, NV, OP, MK, MB, ZA, LG

Visualization: ZA, LG, AC

Supervision: LG, JLM, RL

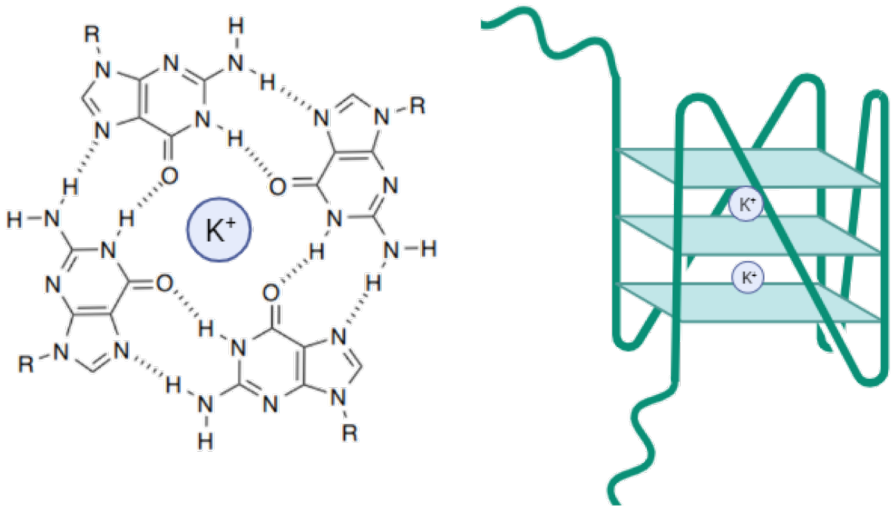
Writing—original draft: ZA, LG

Writing—review & editing: ZA, LG, JLM, VB, RL

Competing interests: The authors declare that they have no competing interests. Data and materials availability: All data needed to evaluate the conclusions in the paper are present in the paper and/or the Supplementary Materials.

Figure 1

A



B

Name	Length	GC %	Number of PQS	PQS frequency
pHV1	85092	55,5	68	0,80
pHV2	6359	56,1	10	1,57
pHV3	437906	65,6	605	1,38
pHV4	635786	61,7	723	1,14
Chromosome	2847757	66,6	4406	1,55

Figure 2

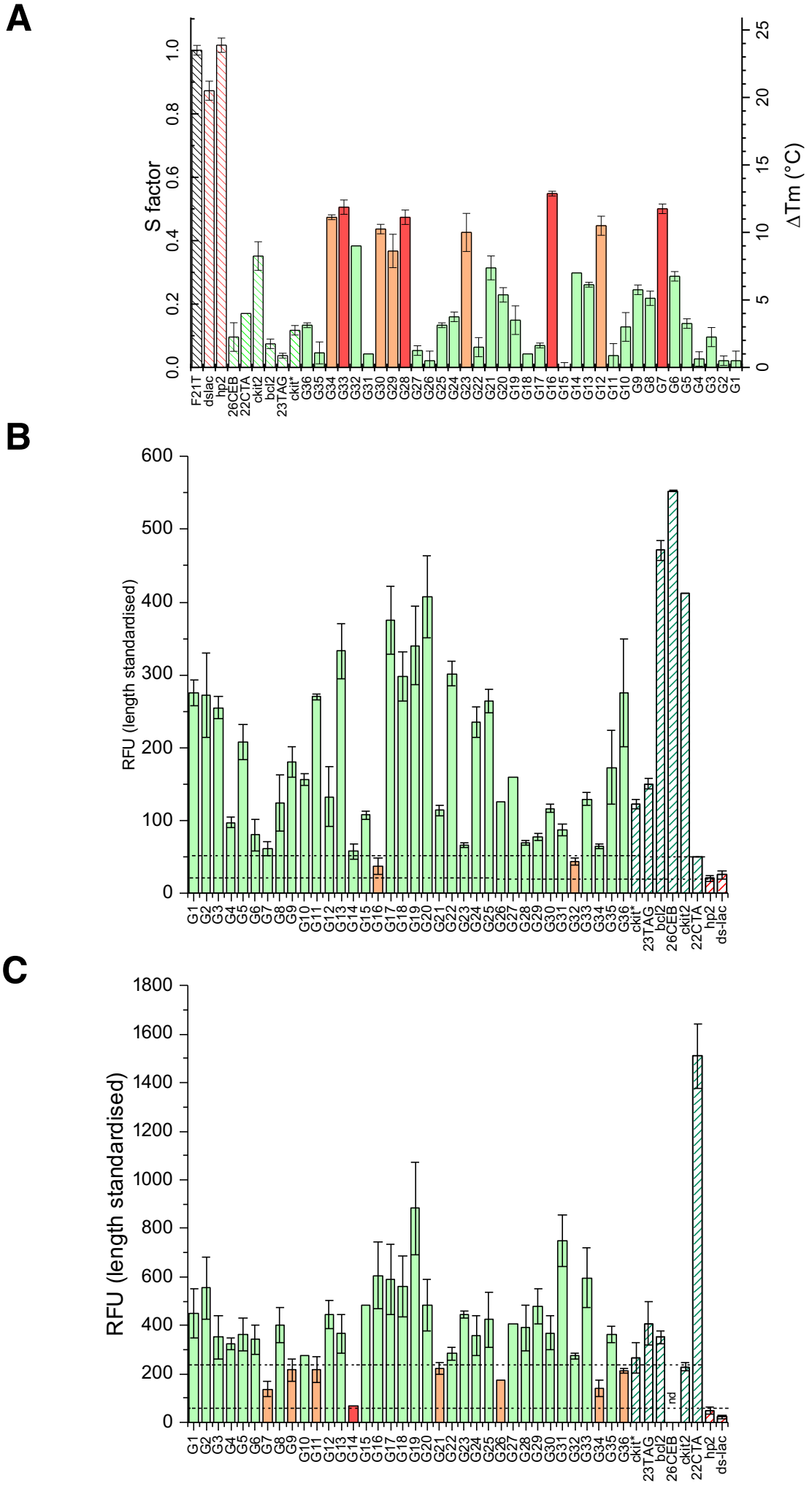

A
S factor
ΔTm (°C)
B
RFU (length standardised)
C
RFU (length standardised)
nd

Figure 3

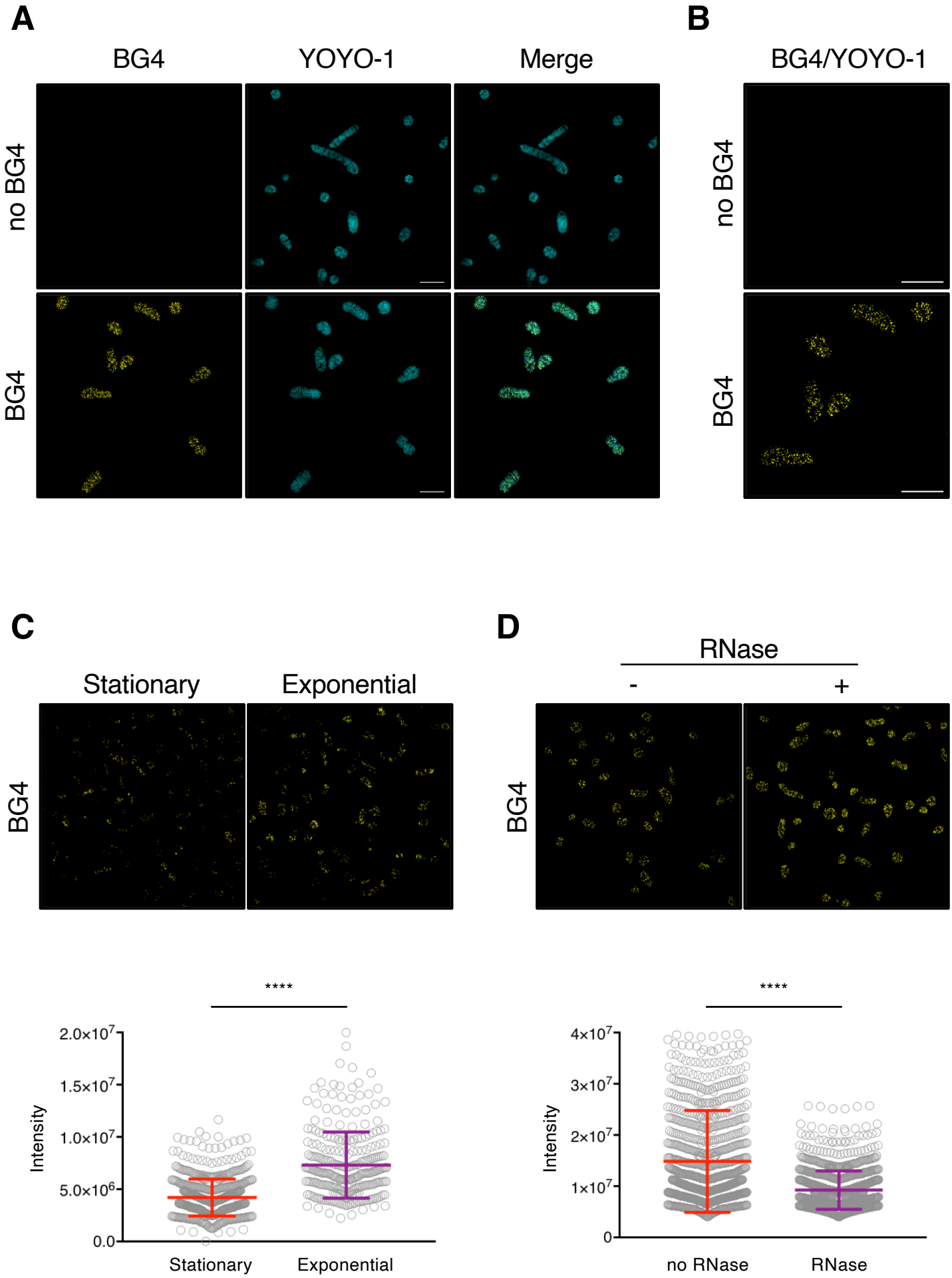
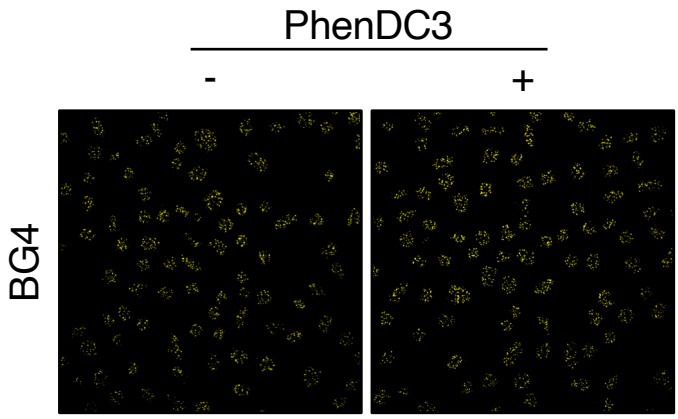
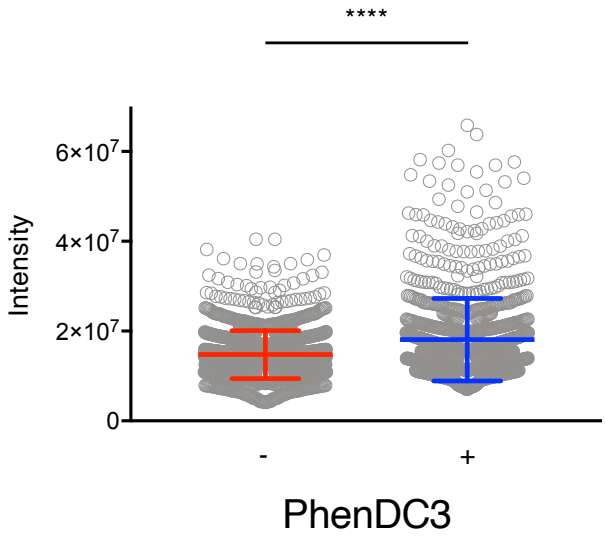

A
BG4
YOYO-1
Merge
no BG4
BG4
B
BG4/YOYO-1
no BG4
BG4
C
Stationary
Exponential
BG4
****
2.0×10^7
1.5×10^7
1.0×10^7
5.0×10^6
0.0
Intensity
Stationary
Exponential
D
RNase
-
+
BG4
****
4×10^7
3×10^7
2×10^7
1×10^7
0
Intensity
no RNase
RNase

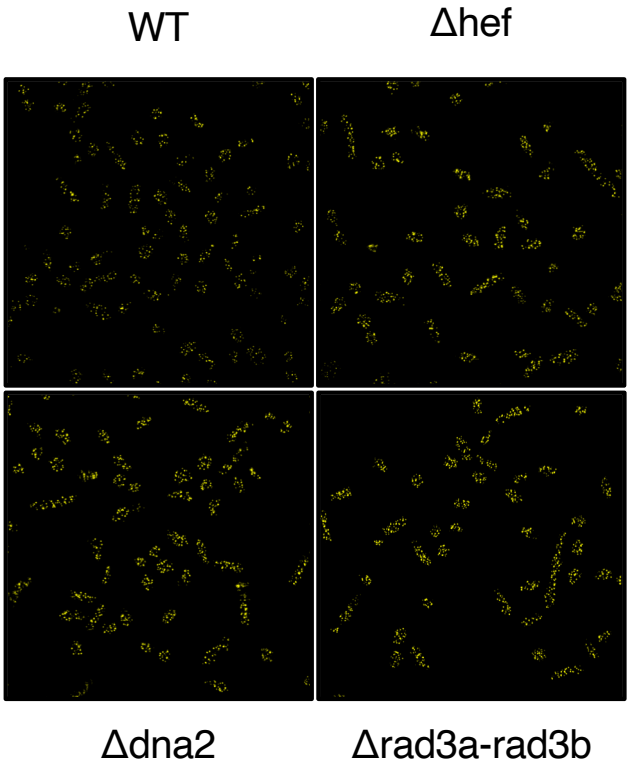
Figure 4

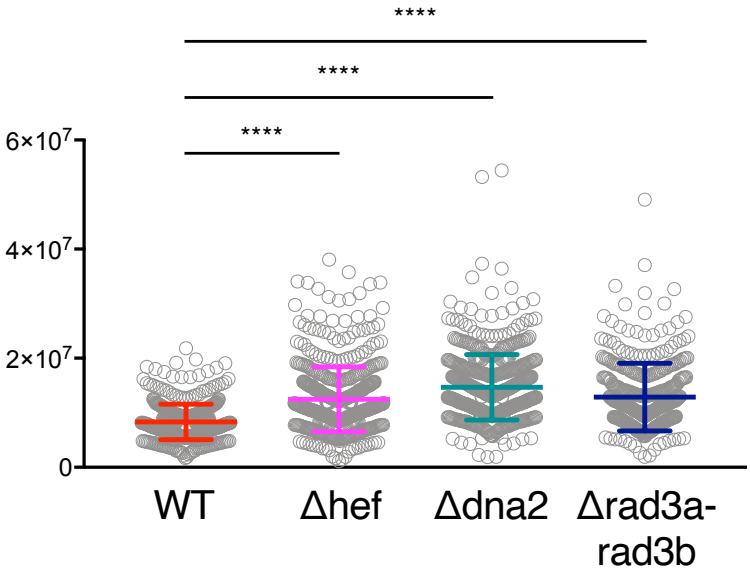
A


PhenDC3
-
+
BG4


****
Intensity
6×10^7
4×10^7
2×10^7
0
-
+
PhenDC3

B


WT
Δhef
Δdna2
Δrad3a-rad3b


****
****
****
6×10^7
4×10^7
2×10^7
0
WT
Δhef
Δdna2
Δrad3a-rad3b