In vivo study of DNA topology in hyperthermophilic archaeon Thermococcus kodakarensis

Paul Villain

université PARIS-SACLAY

In vivo study of DNA topology in hyperthermophilic archaeon *Thermococcus kodakarensis*

Étude in vivo de la topologie de l'ADN chez l'archée hyperthermophile Thermococcus kodakarensis

Thèse de doctorat de l'université Paris-Saclay

École doctorale n°577: structure et dynamique des systèmes vivants (SDSV)
Spécialité de doctorat : Sciences de la Vie et de la Santé

Unité de recherche : Université Paris-Saclay, CEA, CNRS, Institute for Integrative Biology of the Cell (I2BC), 91198, Gif-sur-Yvette, France.
Référent : Faculté des sciences d'Orsay

Thèse présentée et soutenue à Paris-Saclay, le 19/11/2021, par

Paul VILLAIN

Composition du Jury

Stéphanie BURY-MONE Professeure, Université Paris-Saclay	Présidente
Laura BARANELLO Assistant Professor, Karolinska Institutet	Rapporteur / Examinatrice
Hannu MYLLYKALLIO Directeur de Recherche, Ecole Polytechnique	Rapporteur / Examinateur
Didier FLAMENT Chargé de Recherche, IFREMER Brest	Examinateur
Olivier ESPELI Directeur de Recherche, Collège de France	Examinateur

Direction de la thèse

Tamara BASTA Maître de Conférence, Université Paris-Saclay	Directrice de thèse
Ryan CATCHPOLE Postdoctoral Associate, University of Georgia	Co-Directeur

Thèse de doctorat

NNT : 2021UPASL084

Acknowledgements

Un grand merci à Tamara pour la qualité de son encadrement. Tu m'as toujours traité comme un collègue et pas comme un subordonné et je sais que déjà rien que pour cela j'ai été plus chanceux que beaucoup de doctorants. Travailler avec quelqu'un d'aussi gentil, rigoureux et investi que toi a été un réel plaisir et m'a permis de progresser énormément d'un point de vue scientifique.

Un grand merci à Ryan d'avoir accepté à côté de ses projets de participer à mon encadrement. En plus d'être extrêmement sympathique tu es l'un des meilleurs scientifiques que je connaisse et je sais à quel point j'ai été chanceux de bénéficier de ton expérience sur Thermococcus.

Je remercie Stéphanie Bury-Moné, Laura Baranello, Hannu Myllykallio, Didier Flament et Olivier Espéli d'avoir accepté de faire partie du jury qui va évaluer mon travail de thèse.

Je remercie Fedor Kouzine pour ses conseils qui m'ont permis de mettre au point l'utilisation du TMP chez Thermococcus.

Je remercie ma tante, Joëlle, de m'avoir transmis le gout des sciences et fait découvrir le monde de la recherche. Je ne serais sans doute pas devenu biologiste sans ton exemple.

Je remercie mon frère d'avoir pris le temps de m'aider pour analyser mes données sur R. Avoir un frère bioinformaticien c'est un peu comme gagné au Loto en biologie !

Je remercie Augustin pour ses coups de main fréquents. Nos séances de bronzage sous UV vont me manquer !

Je remercie Violette de m'avoir initié avec bonne humeur à la construction d'arbres phylogénétiques.

Je remercie Patrick de m'avoir lancé sur ce projet finalement pas si fou que ça !

Je remercie Florence, Jacques, Marie-Claire et Benjamin d'avoir été des collègues très agréables. L'ambiance du laboratoire va me manquer !

Je remercie Myriam et Corinne d'avoir fait tourner le labo durant toute ma thèse. Quatre ans de thèse ça fait beaucoup de bassines de verrerie…

Je remercie Danièle, Catherine et Tomio pour leur gentillesse. Votre départ a fait un vide dans le labo.

Je remercie Olivier Namy, Isabelle Hatin et Pauline François pour leurs conseils avisés sur le RNAseq.

Je remercie tous les collègues du bâtiment 409 : le groupe de Pascale Servant, l'équipe de Nicolas Bayan, la start-up, Djinthana et Luis, qui m'ont toujours aidé lorsque j'en avais besoin.

J'ai une pensée pour Michael Dubow qui nous a quitté récemment. Tu as été un professeur atypique et extrêmement gentil qui m'a laissé de bons souvenirs.

Je remercie Romain le Bars d'être resté avec moi jusqu'à pas d'heure pour faire toutes ces images de microscopie.

Je remercie toute la plateforme de séquençage, et en particulier Yan, Erwin et Céline pour leur disponibilité et leur compétence.

Je remercie Pauline qui m'a supporté et accompagné pendant ces années de travail intense et de stress.

Je remercie toute ma famille et mes parents en particulier, qui m'ont encouragé tout au long de mes études. Ces week-ends avec vous m'ont permis de décompresser.

Table of contents

Introduction 8

I. Archaea: prokaryotes with eukaryotic flavour 8

A. Discovery of Archaea 8

1. Molecular phylogeny: a breakthrough for understanding Evolution and diversity of organisms 10

2. Specific features of Archaea 13

B. Diversity and Ecology of Archaea 18

1. Inhabitants of extreme environments 18

2. Your discrete neighbours 20

C. Evolution of Archaea 24

1. Phylogenetic tree of Archaea 24

2. The position of Archaea in the tree of life and Eukaryogenesis 24

II. Three-dimensional structure of DNA and chromosomes 27

A. Spatial organisation of chromosomes 27

1. Chromosomal domains 30

2. Topological domains or chromosomal interaction domains 32

3. DNA loops 34

4. Chromosome architects 36

B. DNA supercoiling 42

1. Opening of the double helix creates topological constraints 42

2. Topological parameters 44

3. Physiological consequences of DNA supercoiling 46

C. Regulation of DNA supercoiling 48

1. Topoisomerases 48

2. Supercoiling homeostasis 51

III. Gyrase and reverse gyrase: two atypical topoisomerases 54

A. DNA gyrase 55

1. An ubiquitous and critical enzyme in Bacteria 55

2. Quick and massive gene regulation in response to environmental changes 59

3. Gyrase in Archaea 61

B. Reverse gyrase 63

1. An atypical topoisomerase 63
2. The only hallmark of hyperthermophily 65
3. Uncertain *in vivo* activities 67
IV. Investigating DNA topology in Archaea: the special case of Gyrase 68
1. Gyrase evolution in Archaea 69
2. Gyrase and adaptation to mesophilic lifestyle 70
V. Objectives of the thesis and experimental strategies 71
1. Reconstruct the evolutionary history of gyrase in Archaea 71
2. Describe the consequences of gyrase introduction in a naïve hyperthermophilic archaeon 72
3. Investigate the importance of DNA supercoiling for adaptation to hyperthermophilic lifestyle in archaea 74
Article I 76
Article II 124
Article III 162
Article IV 178
Conclusions and Perspectives 194
I. Evolutive history of gyrase in Archaea 194
II. Consequences of gyrase acquisition by a naïve archaeon 196
III. Importance of DNA supercoiling for adaptation to hyperthermophilic lifestyle 197
IV. Mapping DNA supercoiling along chromosomes by psoralen crosslinking 199
References 202

List of figures

Figure 1: Darwin first draw of phylogenetic tree 9
Figure 2: Nucleic acids can be used to make phylogeny 12
Figure 3: Comparison of lipids from Archaea and Bacteria 14
Figure 4: Methanogenesis in Archaea 15
Figure 5: RNA polymerase from Archaea and Eukaryotes are orthologous 17
Figure 6: Many lineages of archaea are extremophiles 19
Figure 7: Archaea are part of the human microbiome 21
Figure 8: Up to date diversity and phylogeny of Archaea 23
Figure 9: Three-domain and two-domain trees of life imply distinct scenario of eukaryogenesis 25
Figure 10: Compaction of Escherichia coli chromosome in the nucleoid 28
Figure 11: DNA is highly compacted in eukaryotic nucleus 29
Figure 12: Chromosome organization in Sulfolobus 32
Figure 13: Chromosomes are hierarchically organized in bacteria and eukaryotes 33
Figure 14: Chromatin loops and transcription regulation 35
Figure 15: Archaeal histones are very similar to the eukaryotic ones 38
Figure 16: Structure of archaeal hypernucleosome 40
Figure 17: Distribution of NAPs and histones in Archaea 41
Figure 18: Topological consequences of transcription and replication 43
Figure 19: Description of DNA topology parameters 45
Figure 20: Reactions catalysed by topoisomerases 49
Figure 21: Model of supercoiling homeostasis in Bacteria 53
Figure 22: Model of supercoiling homeostasis in Sulfolobus 54
Figure 23: Primary domain structure of E. coli DNA gyrase 56
Figure 24: Molecular model of DNA Gyrase 57
Figure 25: DNA gyrase supercoiling mechanism 58
Figure 26: Chromosomal map of E. coli supercoiling sensitive genes 60
Figure 27: Structural model of Thermotoga maritima reverse gyrase 64
Figure 28: Phylogenetic tree of bacterial and archaeal reverse gyrase sequences 66

List of tables

Table 1: Comparison of the characteristics from Bacteria, Archaea and Eukaryotes 13
Table 2: Main characteristics of known nucleoid-associated proteins 37
Table 3: Main characteristics of known topoisomerases 50
Table 4: Distribution of gyrase in Archaea 62

Abbreviations

3C: Chromosome conformation capture

ADP: adenosine diphosphate

aIF: archaeal translation initiation factor

Alba: acetylation lowers binding affinity protein

ATP: adenosine triphosphate

bp: base pairs

BRE: B recognition element

CIDs: chromosome interaction domains

CTCF: CCCTC-binding factor

CTD: C-terminal domain

DNA: deoxyribonucleic acid

DPANN: Diapherotrites, Parvarchaeota, Aenigmarchaeota, Nanoarchaeota and Nanohaloarchaeota superphylum

Dps: DNA-binding protein from starved cells

eIF: eukaryotic translation initiation factor

ESCRT: endosomal sorting complexes required for transport

ESP: eukaryote signature proteins

FIS: factor for inversion stimulation

FISH: fluorescence in situ hybridization

GTDB: Genome Taxonomy Database

HGT: horizontal gene transfer

H-NS: histone-like nucleoid structuring protein

Hsp: heat shock protein

HU: heat-unstable protein

IHF: integration host factor

kb: kilobase

kDa: kilodaltons

LACA: Last Archaeal Common Ancestor

LBA: long branch artefact

LBCA: Last Bcterial Common Ancestor

LECA: Last Eukaryotic Common Ancestor

Lk: linking number

Lrp: leucine-responsive regulatory protein

LUCA: Last Universal Common Ancestor

Mb: megabase

Mcr methyl-coenzyme M reductase

MMS: methyl methanesulfonate

MNase: micrococcal nuclease

MNase-seq: micrococcal nuclease coupled with deep sequencing

MPa: megapascal

NAPs: nucleoid-associated proteins

NGS: Next Generation Sequencing

pH: potential of hydrogen

Pol II: ARN polymerase II

RNA: ribonucleic acid

RNAP: RNA polymerase

rRNA: ribosomal RNA

SMC: structural maintenance of chromosome complexes

TACK: Thaumarchaeota, Aigarchaeota, Crenarchaeota, Korarchaeota superphylum

TADs: topologically associating domains

TBP: TATA binding protein

TF: transcription factor

TrmBL2: transcription regulator of the maltose-like system 2

Tw: twist

WHD: winged-helix domain

Wr: writhe

Introduction

I. Archaea: prokaryotes with eukaryotic flavour

A. Discovery of Archaea

The description and classification of living organisms plays a fundamental and historical role in Biology. This discipline reaches a turning point in the 18th century with the binomial classification of Carl Von Linné. According to this classification, each organism's name is composed of the general name, the genus, which is shared with several organisms and a specific name, which refers to a particular species. This binomial nomenclature considerably facilitated exchange between scientists and is still used today. It paved the way to the new scientific discipline the Taxonomy, which aims to establish a general classification of living organisms. At the end of the 18th century, Jean-Baptiste de Lamarck introduced the concept of Evolution to explain the diversification of living organisms and their adaptation to their environment. This theory was further developed by Alfred Russel Wallace and Charles Darwin who independently introduced the concept of Natural Selection (Darwin 1859). They postulated that organisms are evolving by chance mutation and those having the most competitive characters are more likely to transmit them to their offspring. To illustrate his theory, Charles Darwin was the first to draw a phylogenetic tree (**Figure 1**) thus setting the stage for the appearance of the new scientific field: the molecular phylogeny.

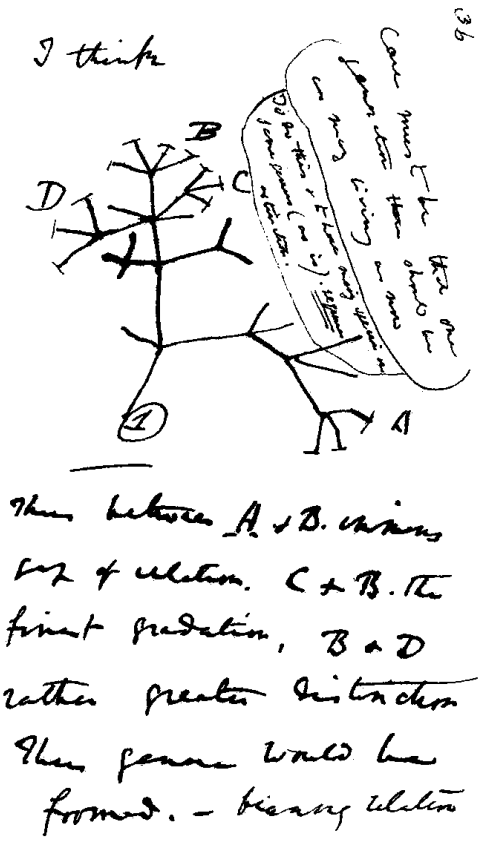


Figure 1: Darwin first draw of phylogenetic tree

"I think case must be that one generation should have as many living as now. To do this and to have as many species in same genus (as is) requires extinction. Thus between A + B the immense gap of relation. C + B the finest gradation. B+D rather greater distinction. Thus genera would be formed. Bearing relation to ancient types with several extinct forms". These notes from Charles Darwin state principles of phylogeny. Parental relations between living organisms can be represented by a tree where each branch represent a distinct lineage, connected by nodes that represent common ancestors. The distance between the end of the branches is proportional to the evolutive distance between the organisms they symbolized. Thus, distantly related lineages like A and B are very far in the tree while close cousins like B and C are just nearby. New lineages are continuously emerging from the tree by speciation at the same time that others get extinct (Darwin 1859).

1. Molecular phylogeny: a breakthrough for understanding Evolution and diversity of organisms

Phylogeny assumes that as long as organisms are sharing a common origin and common features, it is theoretically possible to establish their family relationships. The more similar they are, the more closely related they are (**Figure 1**). This implies that the features used to establish the phylogeny were vertically inherited. Historically, first phylogenies were done by comparing easily describable macroscopic features such as morphological characteristics, thus excluding the use of molecular characters for classification of microorganisms. This limitation disappeared with the work of Emile Zuckerkandl and Linus Pauling who showed that *molecular* characteristics (digestion patterns of proteins) can be used to establish evolutionary relationship between organisms (Zuckerkandl and Pauling 1965). To be useful for phylogenetic analysis, Zuckerkandl and Pauling postulated that the biomolecules must behave as molecular clocks, *i.e.* have a constant rate of mutation over long periods of time. By using fossil record data, it is even possible to calibrate the molecular clocks thus allowing to associate a quantity of time to each branch of the tree.

The concept of molecular clock was a major breakthrough that led to the most famous discovery of modern evolutionary biology (Cavicchioli 2011; Albers et al. 2013). In 1977, Carl Woese and George Edward Fox used the universally conserved small rRNA subunit (16S rRNA) to establish a first biologically meaningful classification of microorganisms based on molecular phylogeny (Fox et al. 1977). By comparing RNase T1 digestion patterns they noticed that methanogenic "bacteria" were distinct from other Bacteria (**Figure 2**). In fact, these data showed that methanogenic "bacteria" were phylogenetically as distant from bacteria, as bacteria were from eukaryotes. They proposed the same year in another article to reclassify these organisms as a third domain of life: the Archaebacteria (Woese and Fox 1977) that was renamed Archaea a few years later to avoid any confusion with Bacteria (Woese, Kandler, and Wheelis 1990). The existence of Archaea as a domain of life distinct from bacteria was supported

by comparisons of complex cellular machineries (see I A 2. “Specific features of Archaea” for more details) such as the RNA polymerase (RNAP). Indeed, the archaeal RNAP is very different from the bacterial one in terms of subunits composition (Zillig 1979).

The work of Carl Woese deeply changed our understanding of the diversity and evolution of living beings, placing for the first time microorganisms in the tree of life. It also helped to promote the evolutionary point of view in all biology fields. Moreover, C. Woese’s classification approach has been successfully adapted to investigate uncultivable microbial diversity by studying small subunit rRNA amplified directly from the environment. It shed light on the fantastic diversity of microorganisms and their strong contribution to geochemical cycles.

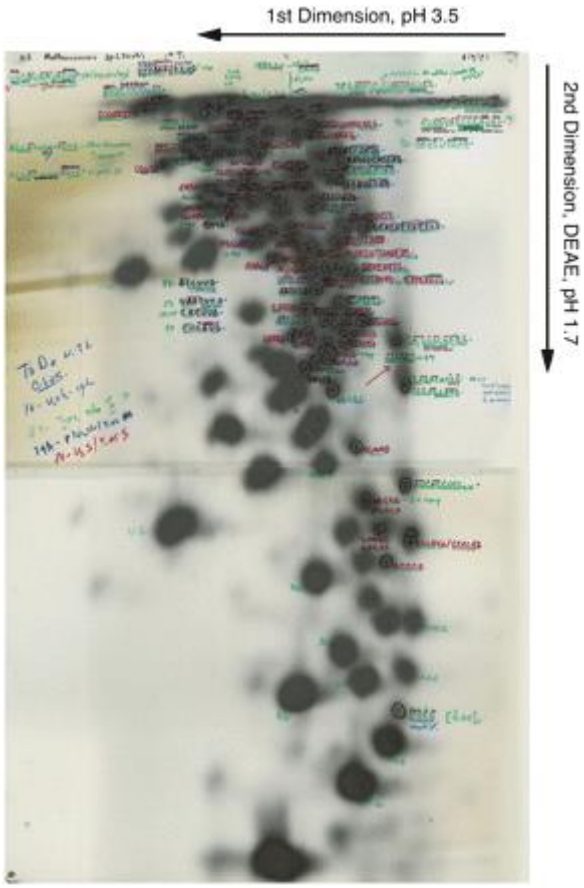


Figure 2: Nucleic acids can be used to make phylogeny

Autoradiography of a T1 ribonuclease digestion of small rRNA subunits resolved by two-dimensional electrophoresis (Pace et al., 2012). Annotations were written by Carl Woese in the course of the discovery of Archaea. The digestion pattern is depending on the sequence of the molecule considered and the accumulation of mutations is proportional to the evolutive distance between lineages. Thus, the comparison of digestion patterns between organisms gives an estimation of their evolutive divergence. But to follow this rule, biomolecules have to respect some conditions towards the organism dataset: (i) they have to be distributed in all organisms considered ; (ii) to have the same function in all of them ; (iii) to be enough informative (i.e. to have residues conserved and enough residues divergent to be compared) ; (iv) to evolved at a similar rate all along the dataset ; (v) to have an evolution rate adapted to the degree of divergence of the dataset (i.e. to evolved quick enough to resolved closely related organisms or slow enough to be informative on distantly related organisms). Biomolecules meeting these conditions are called molecular clocks and can be used to draw phylogenetic trees that reflect the relative degree of relationships of their owners.

2. Specific features of Archaea

After initial discovery of Archaea by Carl Woese, scientists soon realised that these organisms had some particular features in common thus corroborating their classification in a separate domain (**Table 1**). One such feature is the chemical composition of the lipid membranes surrounding archaeal cells. Archaea harbour original lipids that were initially discovered in the extreme halophilic Halobacteria (Sehgal, Kates, and Gibbons 1962; Marshall and Brown 1968). Instead of classical fatty acid ester-linked lipids with a glycerol-3-phosphate backbone, Archaea have isoprenoid ether-linked lipids with a glycerol-1-phosphate backbone. Some archaea from Euryarchaeota and Crenarchaeota phyla also harbour tetraether lipids that span the membrane and form a monolayer that contrast with the classical lipid bilayer of Bacteria and Eukarya (Langworthy, Mayberry, and Smith 1974; De Rosa et al. 1980) (**Figure 3**). All these lipid specificities have deep impact on membrane function, making them less permeable and more resistant to harsh conditions like acid or high temperature.

Table 1: Comparison of the characteristics from Bacteria, Archaea and Eukaryotes

Trait	Bacteria	Archaea	Eukarya
Carbon linkage of lipids	Ester	Ether	Ester
Phosphate backbone of lipids	Glycerol-3-phosphate	Glycerol-1-phosphate	Glycerol-3-phosphate
Metabolism	Bacterial	Bacterial-like	Eukaryotic
Core transcription apparatus	Bacterial	Eukaryotic-like	Eukaryotic
Translation elongation factors	Bacterial	Eukaryotic-like	Eukaryotic
Nucleus	No	No	Yes
Organelles	No	No	Yes
Methanogenesis	No	Yes	No
Pathogens	Yes	No	Yes

Archaea have a mixture of bacterial and eucaryotic traits. They distinguish themselves by special ether-linked lipids built on glycerol-1-phosphate, the ability to perform methanogenesis and an intriguing absence of pathogenicity (adapted from Cavicchioli 2011).

Methanogenesis is another feature so far restricted exclusively to Archaea, meaning that all the biogenic methane, à strong greenhouse gas, is produced by Archaea. Methanogenesis and its counterpart, methanotrophy, are complex enzymatic pathways (**Figure 4)**. But they involve the same key enzyme, the methyl-coenzyme M reductase (Mcr), making it the marker of methane metabolizing organisms (Evans et al. 2019). Methanogenesis was initially discovered in cultured Euryarchaea (see I.B Diversity and Ecology of Archaea) where at least, eight orders of methanogens exist. More recently, Mcr was also detected in several uncultured lineages like Bathyarchaeota and Verstraetearchaeota from the TACK superphylum, indicating that methanogenesis is a more widely spread metabolic pathway in Archaea than initially thought (Adam et al. 2017).

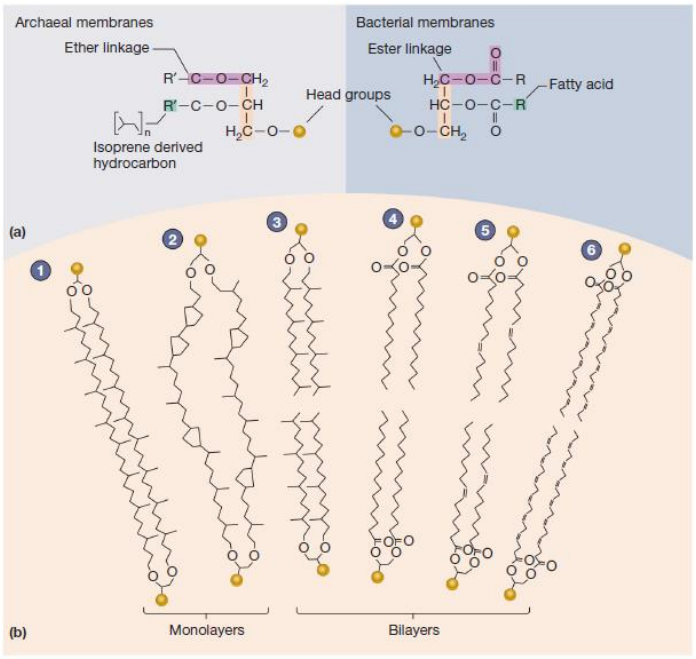


Figure 3: Comparison of lipids from Archaea and Bacteria

(a) Archaeal lipids exhibit an ether linkage connecting their carbon chains to glycerol-1-phosphate, while bacterial and eukaryotic lipids exhibit ester linkage that connect to glycerol-3-phosphate. (b) Lipids 1, 2, 3 are typical of archaea with isoprenoid carbon chains ant the potential to form monolayer. Lipids from bacteria or eukaryotes 4, 5, 6 are built on fatty acid and exclusively form monolayers (adapted from Willey, Sherwood, et Woolverton 2017)

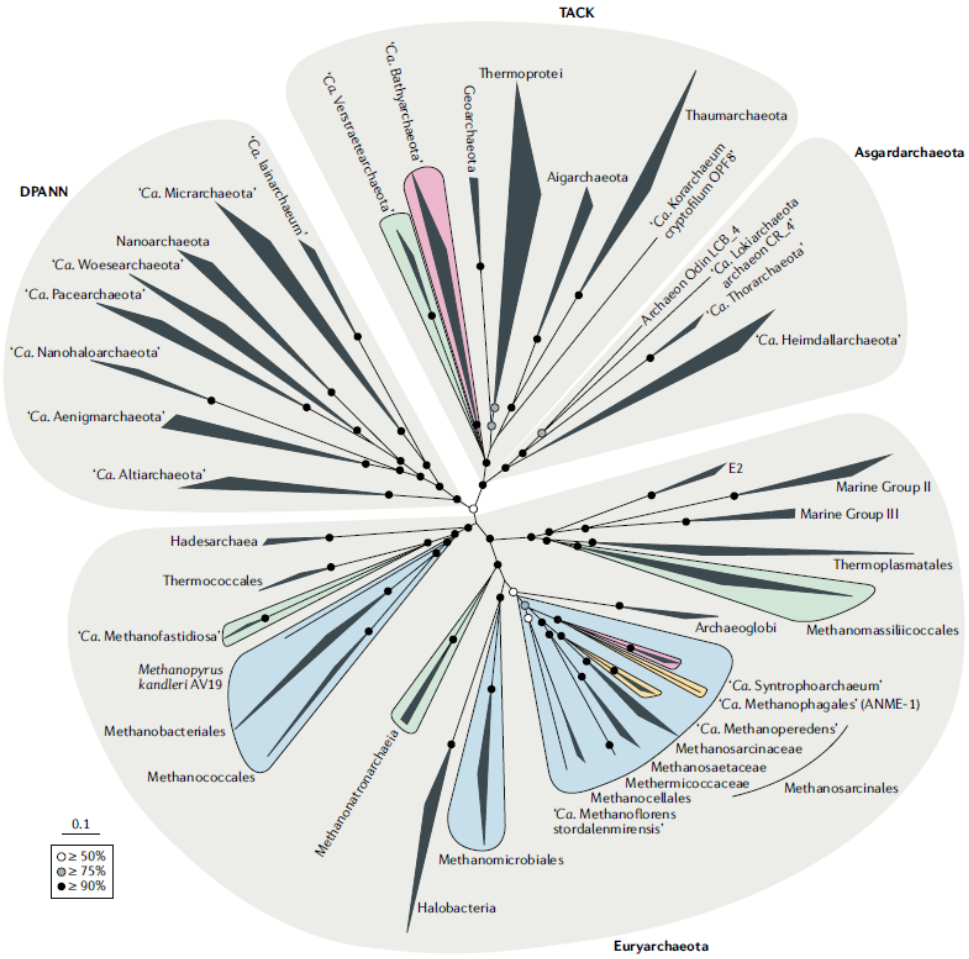


Figure 4: Methanogenesis in Archaea

Phylogeny of Archaea with methanogenic lineages highlighted with colours. The potential for the different lineages to perform methanogenesis is inferred from the presence of *mcr* and *mcr-like* genes in their genome. Most of the methanogenic lineages of archaea are belonging to the Euryarchaeota superphylum (adapted from Evans et al. 2019).

Another specificity of Archaea is the absence of pathogenic species (Cavicchioli et al. 2003). This is intriguing because Archaea are ubiquitous and they interact tightly with a wide range of eukaryotes. A good example are methanogens found in the gut of mammals or associated with plant roots (Vogels, Hoppe, and Stumm 1980; Koskinen et al. 2017; Großkopf, Stubner, and Liesack 1998). Moreover, in humans, besides intestines, the archaea were found in significant proportion also on skin, lungs, nose and oral

cavity (Koskinen et al. 2017). From a molecular point of view, there is no evident explanation for absence of pathogens among archaea. Archaeal genomes encode toxins (Mullis et al. 2019) such as halocins in halophiles and sulfolobicin in Sulfolobales (O'Connor and Shand 2002). They also encode secretion systems (Makarova, Koonin, and Albers 2016) and adhesion systems (Kachlany et al. 2000) which are common virulent factors. And finally, archaeal genomes are dynamic (Wagner et al. 2017), with horizontal gene transfers occurring from Bacteria and Eukarya (HGT), occurrence of different mobile elements and recombination systems (Cossu et al. 2017). Dynamic genomes are crucial for the acquisition of virulence factors and to drive quick, reversible phenotypic changes during the infection process. As Archaea apparently fulfil all the conditions to be pathogens, it is tempting to speculate that we have not discovered enough archaea to find the pathogenic ones. But even if this is true, by comparing the proportion of pathogens in Bacteria and Eukarya with the number of described archaea, there is still an obvious bias against pathogenicity in Archaea (Cavicchioli et al. 2003).

Despite these specificities, Archaea share a lot of characteristics with Bacteria and Eukarya. Like Bacteria, Archaea lack intracellular compartmentation and perform simultaneous transcription and translation. They contain small circular genomes, with a high density of genes that can be organized in operons (Koonin 2009). Along with these typical prokaryotic features, Archaea harbour a number of eukaryotic features. Notably, the factors involved in cellular information processing (transcription, translation, DNA replication and repair) are homologous to eukaryotic equivalents. For example, the transcription in Archaea is performed by a single RNA polymerase (RNAP) homologous to the eukaryotic RNA Pol II (Huet et al. 1983) (**Figure 5**). Both enzymes present a similar crystal structure (Hirata, Klein, and Murakami 2008) and require the same partners (Werner and Grohmann 2011). Archaeal Pol II-like holoenzyme interacts with the eukaryotic-like transcription factors TFB, TFE and TBP, to bind the TATA box and B recognition elements (BRE). Another example is the structure and composition

of ribosome as shown by C. Woese's initial work on small ribosomal subunit. But archaeal ribosome is not the only translation actor with counterparts in Eukarya (Bult et al. 1996; Klenk et al. 1997). Archaeal translation initiation factors (aIFs) aIF-1A, aIF-2, aIF-2B and aIF-5 are all homologous to those of Eukarya (eIFs) (Benelli and Londei 2011) as well as elongation factors (aEF) aEF-1α, aEF-2 and the release factor aRF-1. Moreover, their sequential actions during translation are reminiscent of their eukaryotic homologs.

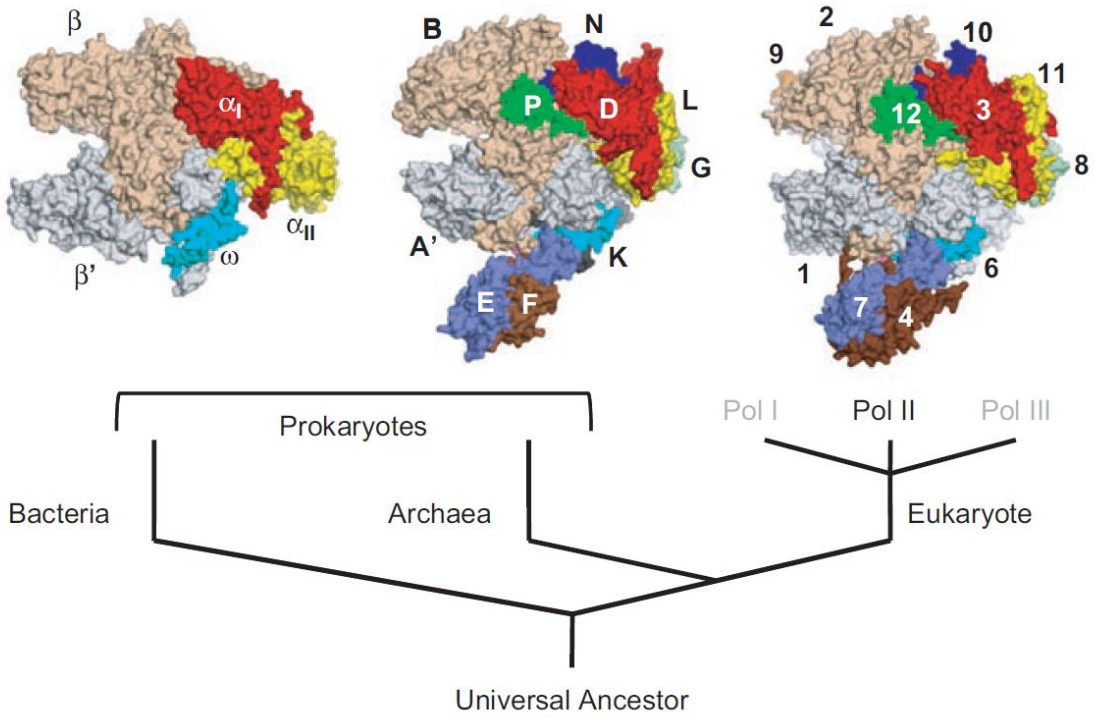


Figure 5: RNA polymerase from Archaea and Eukaryotes are orthologous

Structures of RNA polymerases from *Thermus aquaticus* (bacterium), *Saccharolobus solfataricus* (archaeon) and *Saccharomyces cerevisiae* (eucaryote). Orthologous subunits harbour the same colours (adapted from Jun et al. 2011).

Since the initial recognition of Archaea as a separate domain of life in 1977, much has been learned about their molecular biology. We presently know that these organisms harbour their own specific molecular characteristics but also a curious mix of bacterial

and eukaryotic traits. How these traits have been acquired throughout the history of archaeal lineage and how archaea contributed to the emergence of complex, eukaryotic life forms are fascinating questions for future studies.

B. Diversity and Ecology of Archaea

1. Inhabitants of extreme environments

Traditionally, Archaea have the image of extremophiles. Indeed, most of the initially described archaea where harbouring at least one criterion of extremophily (**Figure 6**). For example, halophilic archaea thrive in highly salty environments like the dead sea, that were thought to be unsuitable for life. This ability to grow in such (from anthropocentric standpoint) a hostile environment stimulated researchers' curiosity and rapidly a lot of extreme halophilic archaea were isolated (Oren 2002), making Halobacteria one of the best described group of archaea.

A lot of archaea are thermophiles or hyperthermophiles, meaning that they grow optimally between 50 and 80°C or above 80°C, respectively. The high diversity of archaea sampled in terrestrial hot springs and hydrothermal vents (Stetter 2006) challenged our vision of life boundaries, showing that even at these temperatures, often in anaerobic or sometimes acidic conditions (Brock et al. 1972; Zillig et al. 1981), life is possible.

To complete the collection, Archaea can also be extreme acidophiles like *Sulfolobus acidocaldarius* that grows optimally at pH 2 (equivalent of acidity in human stomach) or the most acidophilic organism *Picrophilus torridus* that grows optimally at pH 0,7 and that can even grow at negative pH values (Brock et al. 1972; Fütterer et al. 2004) ; extreme psychrophiles like *Methanogenium frigidum* and *Methanococcoides burtonii* that grows optimally at 15°C and 23°C but that can even grow at slightly negative temperatures (Franzmann et al. 1992; 1997) ; extreme piezophiles like *Thermococcus*

barophilus that grows optimally at 40 MPa that is 400 times the atmospheric pressure (Marteinsson et al. 1999) ; extreme alkaliphiles like *Natronomonas pharaonis* that grows up to pH 11 (Soliman and Trüper 1982). But the image of extremophiles that sticks to Archaea is actually misleading, because a lot of them live in much more "reasonable" conditions (DeLong 1998).

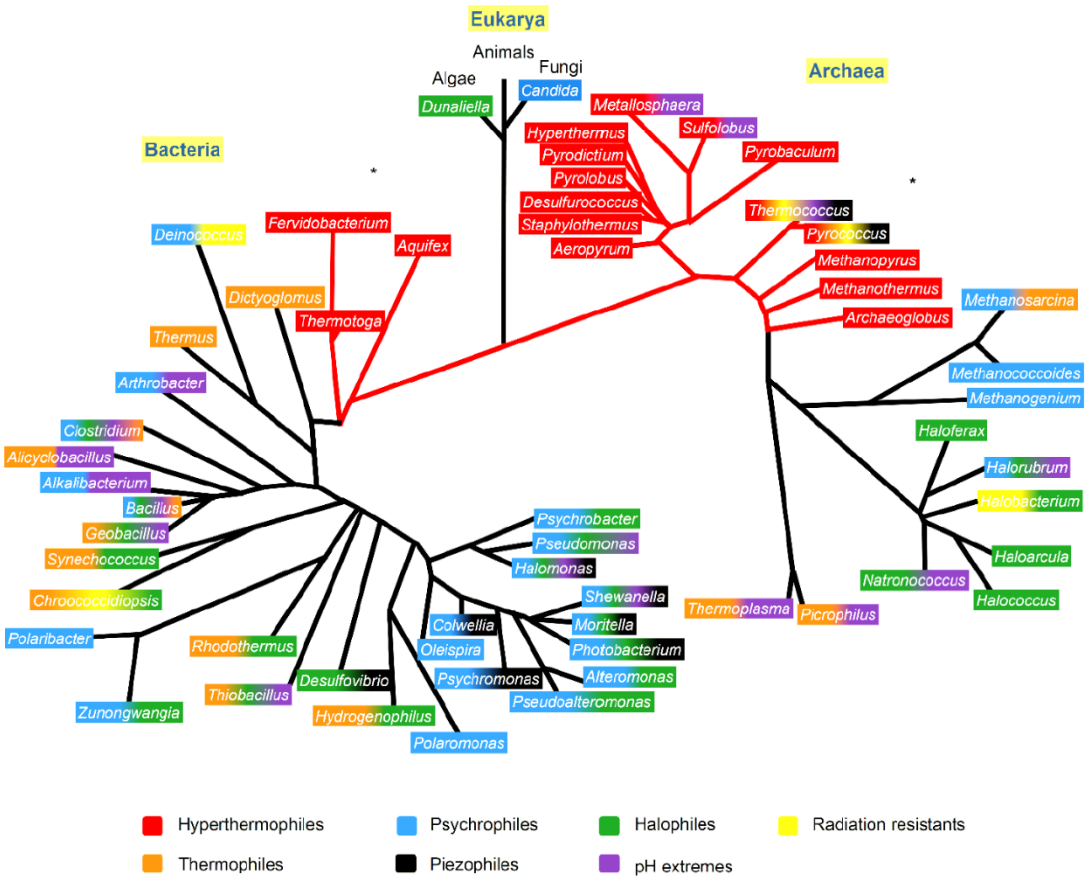


Figure 6: Many lineages of archaea are extremophiles

Phylogeny of archaea coloured according to the type of extremophly present in the different lineages. A lineage is coloured if it encompassed at least one species belonging to the corresponding category of extremophily (adapted from Dalmaso, Ferreira, et Vermelho 2015)

2. Your discrete neighbours

The huge progress in the last 20 years in DNA sequencing technologies completely changed our vision of microbial diversity. Coupled with metagenomic approaches, the microbial diversity can be easily analysed in any environment, with no restriction to cultivable organisms. These analyses revealed that archaea are present in all kinds of environments and not just in extreme ones (Adam et al. 2017). For example, a significant part of ocean microbial diversity is made of archaea (Karner, DeLong, and Karl 2001; Corte et al. 2009). While the proportion of archaea in different samples from marine water is strongly variable, it seems that their numbers increase proportionally with the depth. This indicated that Archaea may be major players in cycling of organic matter in the deep sea (Karner, DeLong, and Karl 2001). A good example of the ubiquity and ecological importance of Archaea is the Thaumarchaeota phylum that was discovered in sea water and initially described as "non-thermophilic Crenarchaeota" (DeLong 1992; Fuhrman, McCallum, and Davis 1992; Brochier-Armanet et al. 2008). Interestingly, these organisms encoded ammonia monooxygenase an enzyme that catalyses the oxidation of ammonia, the first step of biological nitrification, a process which was believed to be exclusive to bacteria. The discovery of these ammonia-oxidising archaea at a high proportion in marine environment and soil, highlighted the fact that Archaea are important contributors to the global nitrogen cycling (Offre, Spang, and Schleper 2013). More recently, Thaumarcheota have also been found to participate significantly to human microbiome. They are found principally on the skin, while Euryarchaeota are dominant in the nose and digestive tract, and Woesearchaeota in lungs (Koskinen et al. 2017; Borrel et al. 2020) (**Figure 7**).

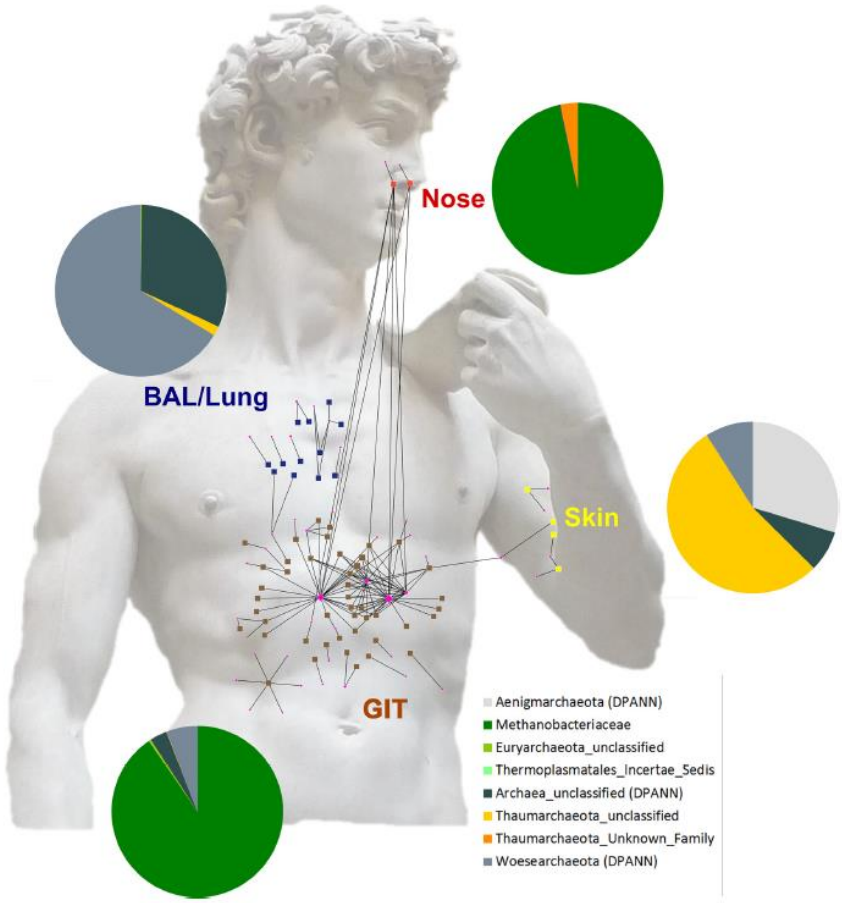


Figure 7: Archaea are part of the human microbiome

Taxonomic distribution of archaea on human body. Archaea were sampled: from lung through bronchoalveolar lavage (BAL) (blue), from skin (yellow), from nose (red), from gastrointestinal tract (GIT) (brown). Similarities between datasets are plotted as a network (adapted from Koskinen et al. 2017).

Numerous other new phyla of Archaea have been discovered from metagenomics studies (**Figure 8**). These new proposed phyla expanded from initially 2 (Crenarchaeota and Euryarchaeota) to 27 the number of major archaeal lineages (Baker et al. 2020). Some of the more recently discovered phyla include Bathyarchaeota methanogens which were recovered from deep aquifers (Evans et al. 2015) and Vestraetearchaeota methanogens from bioreactors (Vanwonterghem et al. 2016). Metagenomic studies have also allowed the surprising discovery of a wealth of nanosized archaea accounting for about a third of the archaeal diversity! These organisms contain atypical genomic

features and thus could not be recovered initially using Archaea-specific 16S rDNA amplification. They were regrouped into the proposed DPANN superphylum, composed of: Diapherotrites, Parvarchaeota, Aenigmarchaeota, Nanohaloarchaeota, Nanoarchaeota (Rinke et al. 2013), Woesearchaeota, Pacearchaeota, Micrarchaeota, Huberarchaeota and potentially, Altiarchaeota (Baker et al. 2020). DPANN were detected in a wide range of environments. They generally have small genomes of about 1 Mbp and they seem to have underwent intensive horizontal gene transfers. These characteristics suggested that DPANN are fast-evolving species with parasitic or symbiotic lifestyle.

Since their discovery by Carl Woese, Archaea have definitively earned a major place in the tree of life over the last 40 years. Their ubiquity, specific metabolisms, eukaryotic features and importance in geochemical cycles make them of great interest. But despite great progresses to sample their diversity and understand their ecology, the principal forces driving this diversification and adaptation of Archaea to a wide range of environments is still an interesting and largely unresolved question.

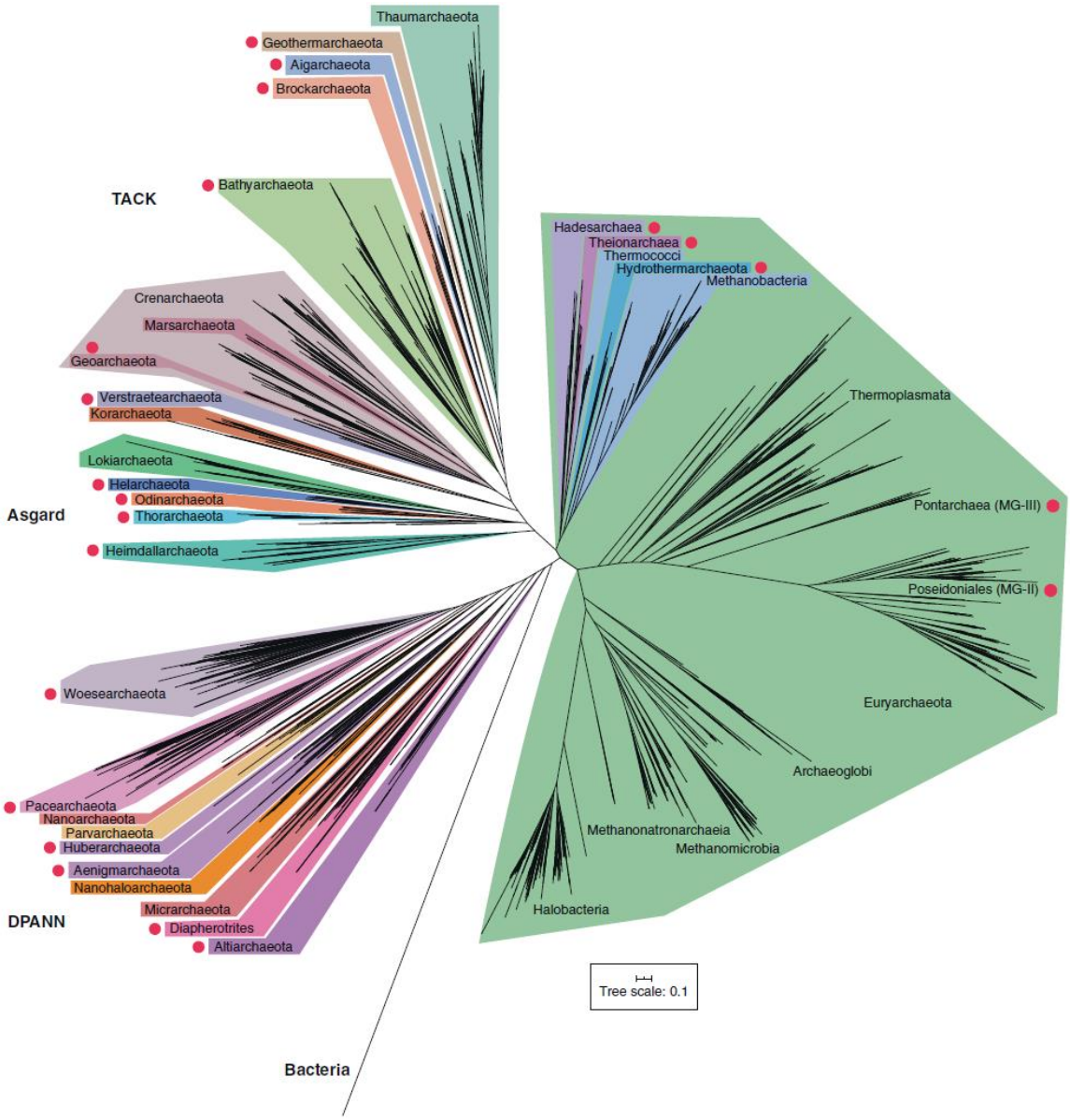


Figure 8: Up to date diversity and phylogeny of Archaea

Metagenomic considerably increased the number of knwon archaea lineages over the last two decades. This recent phylogeny of Archaea was generated using 36 conserved proteins in 3,549 archaeal genomes and 40 bacterial genomes. Red dots indicate the absence of cultured representative in lineages (adapted from Baker et al. 2020).

C. Evolution of Archaea

1. Phylogenetic tree of Archaea

The fast expansion of archaeal lineages has strongly complexified the initial dichotomic archaeal tree (**Figure 8**), divided between Euryarchaeota and Crenarchaeota. Newly discovered uncultured lineages provide lots of information about archaeal ecology, metabolism and evolution but are challenging to place on the phylogenetic tree. For instance, it is still unclear if the monophyly of DPANN is real or just a long branch artefact (LBA) due to their propensity to evolve fast. Indeed, such fast-evolving species accumulate mutations, making them more similar for tree construction algorithms that artificially group them, often at the base of the tree, closest to the outgroup.

The profusion of new archaea is also a taxonomic problem. Taxonomic ranks assigned to archaeal lineages are not necessarily homogeneous in term of genomic divergence. This problem, general to all microorganisms, can be explained by the early phenotypic classifications that often reflect convergence rather than evolutionary relationships and the over-splitting of extensively studied groups irrespectively of their real divergence. The emergence of the new Genome Taxonomy DataBase (GTDB) is a recent attempt to try to standardize and equalize classification all along the tree of life (Parks et al. 2018, 2020). To date four major phyla of archaea are recognize: Euryarchaeota, TACK, DPANN and the recently described Asgards. This last group is particularly important from an evolutionary point of view, as detailed below.

2. The position of Archaea in the tree of life and Eukaryogenesis

The position of the Archaea on the tree of life with respect to Eukaryotes, is of special interest to understand eukaryogenesis, one of the major evolutionary questions. Indeed, two contradictory hypotheses, highly dependent on the archaeal phylogeny, have been proposed. In the first one, Eukarya are sister group of Archaea, giving the

"Woese" three-domains (3D) tree of life. In the second one, Eukarya emerged from within Archaea, giving a two-domains (2D) tree of life where Archaea are not monophyletic. This hypothesis is known as the "Eocyte" one, according to the first group of archaea (current Crenarchaeota) proposed to be at the origin of Eukarya (Lake et al. 1984). The 2D and 3D tree topologies, imply distinct scenario of eukaryogenesis. A 3D tree of life means that complex features of Eukarya, like the spliceosome, the nucleus, the ability to perform meiosis or mitosis, were acquired step by step by a protoeukaryote after the eukaryotic lineage had split from archaeal one (**Figure 9**). Coupled with the crucial endosymbiosis of an alphaproteobacterium to form mitochondria, this complexification would conduct to modern eukarya (Spang, Caceres, and Ettema 2017). Conversely, in the 2D topology, many complex eukaryal features must have emerged relatively fast , in the relatively short period between splitting from archeal lineage and the last common eukaryotic ancestor. This fast evolution period could be explained by the early acquisition of mitochondria, that would dramatically

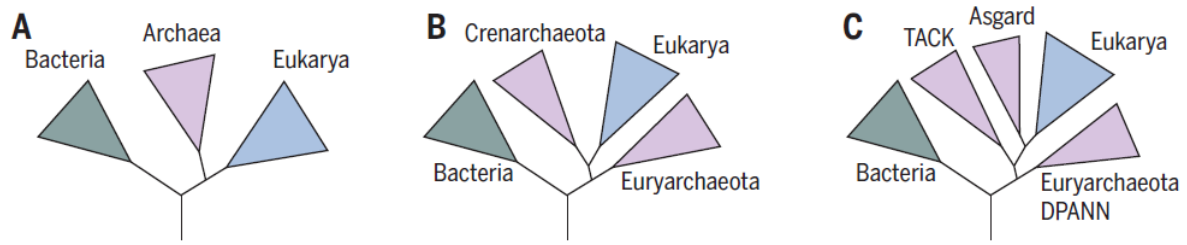


Figure 9: Three-domain and two-domain trees of life imply distinct scenario of eukaryogenesis

Schematic representation of the tree of life with Bacteria, Archaea and Eukaryotes as three distinct domains (A) or Eukaryotes arising from Archaea (B and C) (adapted from Spang, Caceres, et Ettema 2017). A three-domains tree of life implies that Eukaryotes evolved their specific features (intracellular compartmentation, cytoskeleton, organelles) independently of Archaea, while in a two-domains tree of life all these features were evolved from an archaeon probably already very similar to modern archaea. The recent discovery of Asgards archaea, that share characteristics with Eukaryotes, has given support to the two-domains tree of life (C). However, the topology of the tree of life is still matter of intense debates.

relax evolutive constrains by providing the necessary burst of energy (Lane and Martin 2010) thus allowing the development of complex features.

The 2D topology gained in notoriety with the discovery of "eukaryote signature proteins" (ESPs) in some Thaumarchaeota (Ettema, Lindås, and Bernander 2011; Yutin and Koonin 2012; Lindås et al. 2008) and, especially, in the recently discovered Asgardarchaeota superphylum (Spang et al. 2015). These ESPs correspond to proteins thought to be absent from prokaryotes but almost ubiquitous in Eukaryotes. Their massive discovery in Asgards suggests that Eukaryotes can emerge from Archaea instead of forming an individual domain. The first asgard metagenomes were recovered in deep marine sediments from the Arctic Mid-Ocean Ridge, next to the Loki's Castle hydrothermal vent and, accordingly the corresponding archaea were named Lokiarchaeota. The phylogenetic analysis of 36 markers proteins suggested that Asgards formed a monophyletic group with Eukarya that was branching in TACK archaea. Moreover, Lokiarchaeota harboured 29% of bacterial genes and about 3% of eukaryotic genes. Among these genes, a lot of ESPs were present, notably: (i) five actin homologues, (ii) a lot of small GTPases from the Ras superfamily (2% of lokiarchaea genome) known to regulate cytoskeleton dynamics (iii) a cluster of ESCRT genes and other genes involved in multivesicular endosome pathway, suggesting an active vesicle trafficking system. These findings were confirmed later with the Thorarchaeota, Heimdallarchaeota, Odinarchaeota, Helarchaeota asgards metagenomes that enriched the pantheon of Nordic gods (Seitz et al. 2016; Zaremba-Niedzwiedzka et al. 2017; Seitz et al. 2019) and with the isolation and sequencing of the lokiarchaeon *Prometheoarchaeum syntrophicum* (Imachi et al. 2020).

However, the 2D topology is still controversial with plenty different opinions. Recent studies argue against 2D topology because of (i) the presence of fast-evolving species in the dataset; (ii) a subset of universal markers supports the 3D topology; (iii) asgard metagenomes are contaminated (Da Cunha et al. 2017; 2018); (iv) phylogenetic signal is too weak in protein sequences to resolve a universal tree contrary to structural motifs

(Harish and Kurland 2017a; 2017b). Spang *et al.* rejected all these arguments (Spang et al. 2018), as other authors (Williams et al. 2020). From another perspective, it has been claimed that only the placement of the root of the Archaea-Eukarya phylogeny can discriminate between a 2D or a 3D topology (Fournier and Poole 2018). Indeed, the root is essential to orientate the sequence of evolutionary events and thus, discriminate between ancestral and recent characters. Recent studies currently propose different rooting: between Euryarchaeota and TACK, within Euryarchaeota, or between DPANN and other archaea (Petitjean et al. 2014; Raymann, Brochier-Armanet, and Gribaldo 2015; Williams et al. 2017). Further metagenomes and phylogenetic analyses will hopefully allow to distinguish between the different points of view.

II. Three-dimensional structure of DNA and chromosomes

A. Spatial organisation of chromosomes

DNA is the molecule supporting genetic information in living organisms. While fascinating in many respects, maybe the most striking characteristics of DNA are its high levels of condensation and organisation. Two main reasons explain these general features. (i) The important amount of information encoded in genomes imposes long molecules that must fit in small cells. As an example, *Escherichia coli* circular genome is made of approximatively 4.6 x 10^6 bp that would form 1.5 mm long ring if relaxed, condensed in a 500 μm^3 by Brownian motion (Verma, Qian, and Adhya 2019). Knowing that *E. coli* measures on average 1 µm long by 0.25 µm large, resulting approximatively in a 0.05 μm^3 cylinder, it is crucial to compact DNA to fit into the cell and even more by considering that the volume of nucleoid is less than 1 μm^3 (**Figure 10**).

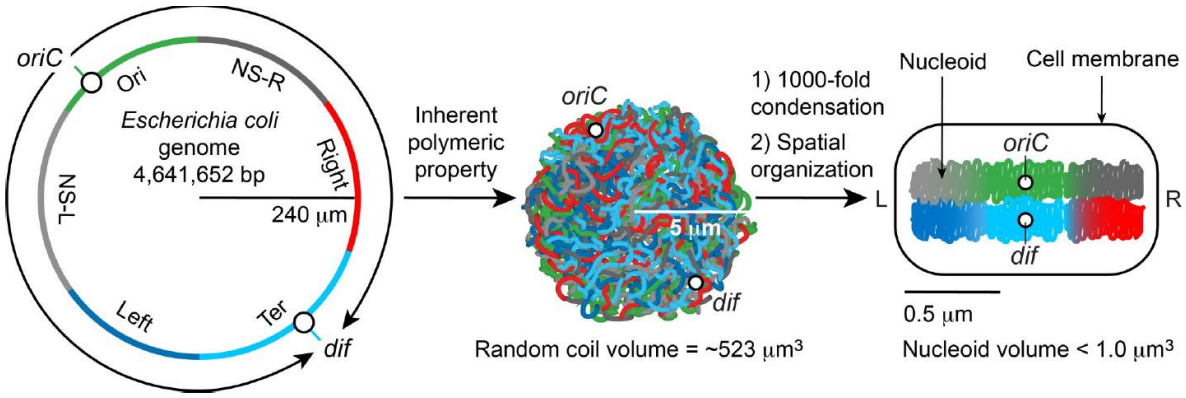


Figure 10: Compaction of Escherichia coli chromosome in the nucleoid

Circular chromosome of *E. coli* occupies a 523 µm3 volume at thermal equilibrium without supercoiling or any stabilizing factors. This random coil must be compacted at least 1000-fold by various chromatin architects to form a nucleoid smaller than 1 μm^3. *E. coli* macrodomains are indicated by colours (adapted from Verma, Qian, et Adhya 2019).

This is also true in the nucleus of Eukaryotes or in every other organism (**Figure 11**). To fit in such a small space, DNA is compacted at least 1000-fold by supercoiling and interactions with various architectural proteins.

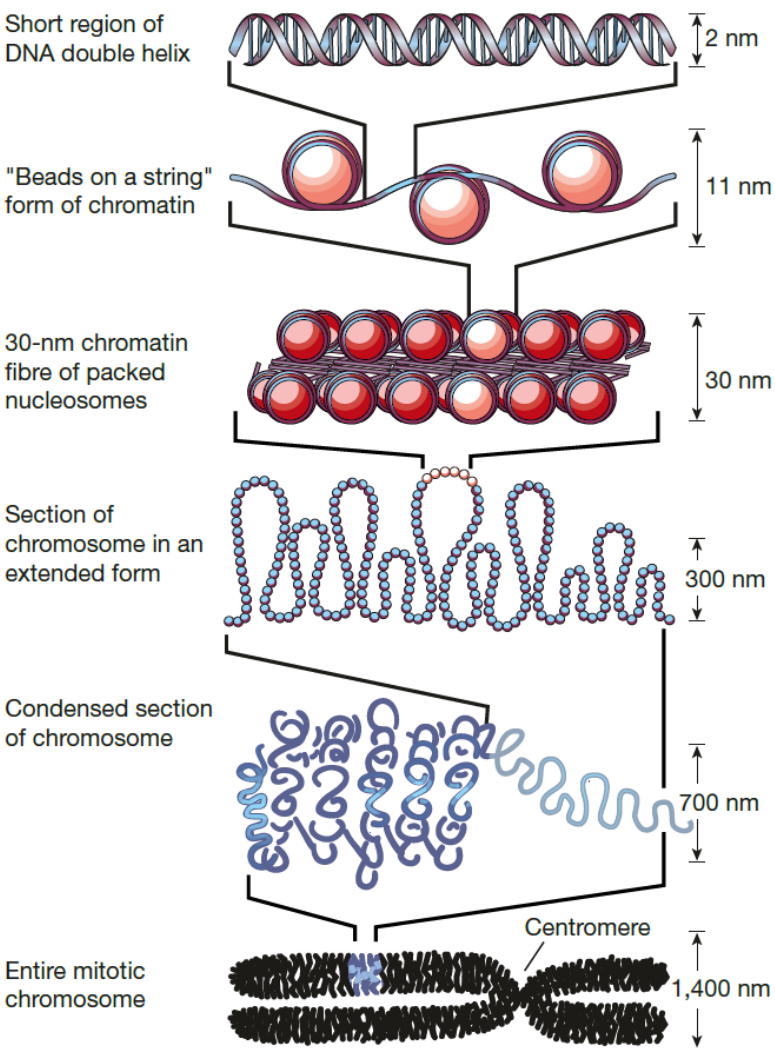


Figure 11: DNA is highly compacted in eukaryotic nucleus

To fit in nucleus, chromosomes from eukaryotes must be compacted more than a hundred-time in length. This compaction involves different levels: first DNA is wrapped around nucleosomes to form the bead on a string structure; then beads are packed together in a 30 nm chromatin fibre; these fibres are folded in higher-order structures by structural proteins leading to highly condensed chromosomes (adapted from Felsenfeld and Groudine 2003).

The genome expression is a tightly regulated process involving numerous proteins and chemical signals. DNA compaction is playing a key role in this process by modulating the accessibility of transcription machinery to different genome regions. A well-known

example is the heterochromatin in Eukaryotes that mostly contains non-transcribed genes (Janssen, Colmenares, and Karpen 2018). Functional compaction was discovered progressively by fluorescence in situ hybridization (FISH) and chromosome conformation capture methods (3C, Hi-C) (Kempfer and Pombo 2019). These findings revealed that the genomes are highly organized at different scales as detailed in the following sections.

1. Chromosomal domains

In Bacteria, chromosome can be divided in large Mb domains, grouping multiple topological domains, that have a higher probability of interaction with each other than with other parts of the genome. This level of genome organization in Bacteria was first brought to light in *Escherichia coli*, thanks to fluorescence in situ hybridization (FISH) showing a co-localization of some genomic markers in the nucleoid (H. Niki, Yamaichi, and Hiraga 2000; Hironori Niki and Hiraga 1998). In this way, two macrodomains were discovered and named Ori and Ter because they encompassed the origin and termination of replication of the chromosome. Two other macrodomains framing the Ter macrodomain, were discovered later on by measuring the rate of recombination of two att sites from the lambda phage inserted at different places in the chromosome (Valens et al. 2004). These four macrodomains are covering all the *E. coli* genome except two non-structured regions, called NS-left and NS-right (**Figure 10**). More recently, chromosome capture technics like 3C or Hi-C have confirmed the previous studies and detailed the structure of macrodomains (Lioy et al. 2018). These technics combine Next Generation Sequencing (NGS) and genomic DNA crosslinking to map each internal contact on the chromosome. Results allowed to precise the borders of macrodomains and, coupled with genetic experiments, to determine actors involved in their formations. It also gave an idea of the range of distance in which interactions occurred, showing that DNA-DNA interactions are restricted to 280 kb in Ter

macrodomain but can reach up to 1 Mb in the rest of the chromosome. While most of the work on bacterial chromosome structure has been performed in *E. coli*, it appears that chromosomes of other distant model bacteria like *Bacillus subtilis*, *Caulobacter crescentus*, *Streptomyces ambofaciens* or *Vibrio cholerae* exhibit similar characteristics (Dame, Rashid, and Grainger 2020; Val et al. 2016).

In Eukaryotes, chromosomes occupy a precise place in the nucleus rather than being all mixed together (Cremer and Cremer 2010). Each volume occupied by a specific chromosome in the nucleus is defined as a chromosome territory, in which, intra-chromosomic interactions are favoured. These territories are divided into two compartments: the A compartment that is enriched in euchromatin and highly transcribed genes and the B compartment that is enriched in heterochromatin and weakly transcribed genes. Euchromatin correspond to an open state of chromatin, accessible to transcription, by opposition to the heterochromatin that is highly condensed. The compartmental organization of chromatin is thought to be advantageous to concentrate transcription machinery components in close proximity to highly expressed genes. Chromatin from the A compartment tends to locate at the periphery of the nucleus while the chromatin from the B compartment is more central.

While chromosome architecture is studied in Bacteria and Eukaryotes since two decades, the question of chromosome architecture in Archaea has been tackled only very recently through two studies based on Hi-C experiments. The first one established a chromosome compartmentalization similar to Eukaryotes in the Crenarchaeota *Sulfolobus acidocaldarius* and *Sulfolobus islandicus* (Takemata, Samson, and Bell 2019). As in Eukaryotes, highly transcribed genes of Sulfolobus and replication origins locate in the A compartment while mobile genetic elements, poorly and conditionally expressed genes are in the B compartment (Takemata and Bell 2020) (**Figure 12**). The second study has been performed in *Haloferax volcanii*, *Halobacterium salinarum* and *T. kodakarensis*, that belong to the Euryarchaeota phylum, and shows a more bacterial-like chromosome organization with no A/B compartmentalization (Cockram et al.

2021). These discrepancies between Euryarchaeota and Crenarchaeota suggest an interesting diversity of chromosome architectures in Archaea, that is waiting to be further explored.

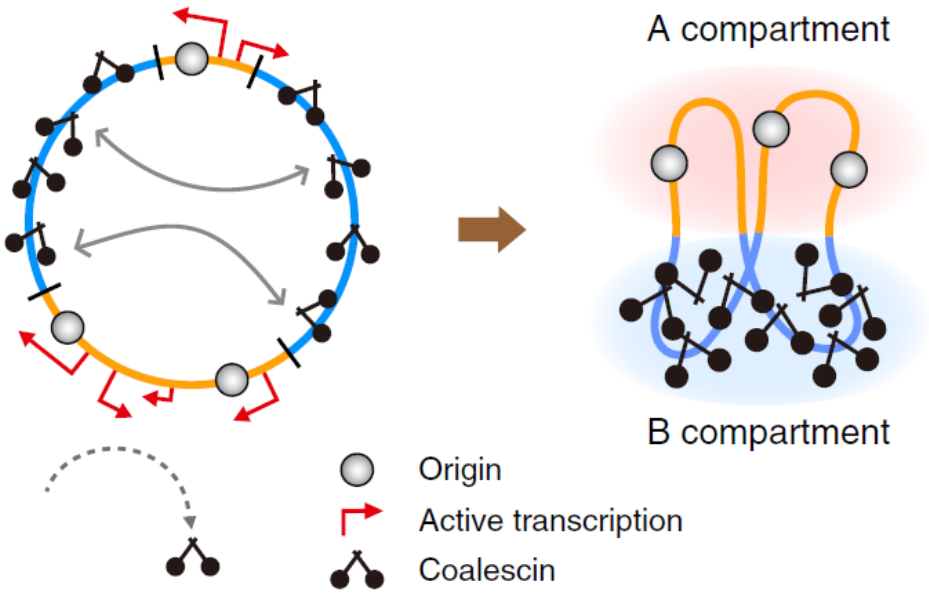


Figure 12: Chromosome organization in Sulfolobus

Schematic representation of *Sulfolobus* archaea chromosome organization (adapted from Takemata, Samson, et Bell 2019). *Sulfolobus* chromosomes are organized into topological domains that interact to form A and B compartments in a reminiscent way of eukaryotes. A compartment concentrates replication origins and highly transcribed genes while B compartment concentrates less expressed genes. The structural maintenance protein (SMC) coalescin, specific of *Sulfolobus*, is involved in the structuration of B repressive compartment.

2. Topological domains or chromosomal interaction domains

To take advantage of the torsional energy stored in DNA, chromosomes supercoiling is tightly regulated and chromosomes are compartmentalized into topological domains (Sinden and Pettijohn 1981). This level of chromosome organization seems to

be common to every organism (**Figure 13**). However, it can differ in terms of domain size and importance according to the species considered. Topological domains are

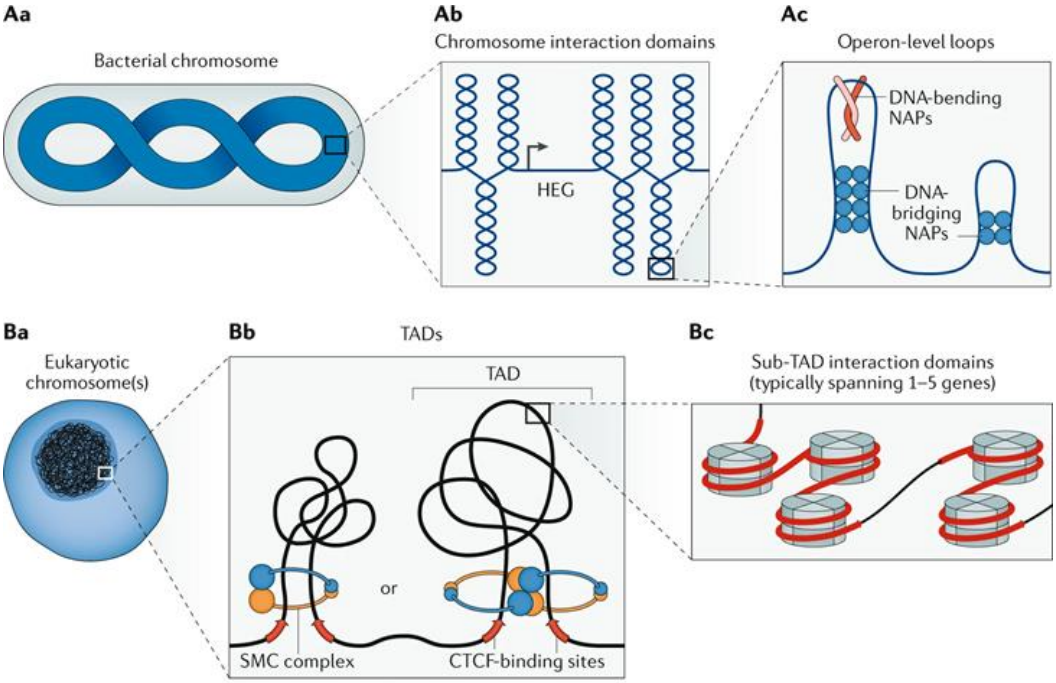


Figure 13: Chromosomes are hierarchically organized in bacteria and eukaryotes

Schematic representation of bacterial and eucaryotic chromosomes architectures (adapted from Dame, Rashid, et Grainger 2020). At the whole scale, bacterial chromosomes are spirally folded in the nucleoid and divided into large macrodomains (not detailed on this schema) (Aa). At a smaller scale (10-100 kb) chromosomes are subdivided into chromosome interaction domains (CIDs) (Ab). CIDs exhibit self-interaction and are insulated from flanking chromatin. Bacterial CIDs are nested, meaning large CIDs can encompass smaller CIDs. At the smallest scale, bacterial CIDs are organized into loops, mostly structured by NAPs, that insulate groups of genes or operons from the rest of the chromosome (Ac).

Chromosomes from eukaryotes are confined into nucleus where they each occupy a defined territory and are split in a highly transcribed A compartment and a less transcribed B compartment (not shown on this schema) (Ba). At a smaller scale, eukaryotic chromosomes are subdivided into topologically associating domains (TADs) (Bb). TADs are formed by loop extrusion through SMC-complex until collision with CTCF-DNA complex occurs. At the smallest scale, sub-TAD domains can be detected corresponding to one to five insulated genes (Bc).

delimited by topological barriers that block the diffusion of supercoils. This allows the coexistence of various topological states in the same chromosome. Topological barriers can be made of combinations of (i) proteins that wrap or loop DNA (histones, NAPs, CTCF, SMC) ; (ii) palindromic repeats (BIMES elements in *E. coli*) ; (iii) actively transcribed regions (Verma, Qian, and Adhya 2019). All these elements cooperate in a complex and dynamic manner, that is not well understood for the moment.

3. DNA loops

Inside topological domains, DNA is organized depending on the prevalence of packaging proteins like histones, NAPs or SMC proteins. In *Escherichia coli* about half of the chromosome is constrained by NAPs (Bliska and Cozzarelli 1987) and the rest is forming plectonemic loops, while nearly all the DNA is constrained by histones in Eukaryotes (Sinden, Carlson, and Pettijohn 1980). This creates various architectures, but a common point is the importance of DNA loops that create a favourable environment for a lot of DNA transactions by promoting DNA contacts. These loops can for instance (i) bring together promoters and transcription terminators, thus promoting RNA-polymerase recycling and increasing transcription rate (**Figure 14a**) (O'Sullivan 2004). DNA loops can also (ii) bring enhancers or silencers in close proximity to a promoter to regulate its transcription (**Figure 14b and 14c**) (Tolhuis 2002 ; Tiwari 2008). In those configurations, enhancer or silencer is bound by a transcriptional regulator that will regulate the transcription initiation complex. Another possibility is (iii) the isolation of a portion of the genome from its genomic context to prevent aberrant regulation. Such structures have been first described in the *Drosophila* flies (Hou 2012). This kind of loops involve insulator elements that are bound by proteins like SMCs or the CTCF in Eukarya. Insulators are present on both sides of the isolated region and form the loop by collapsing together.

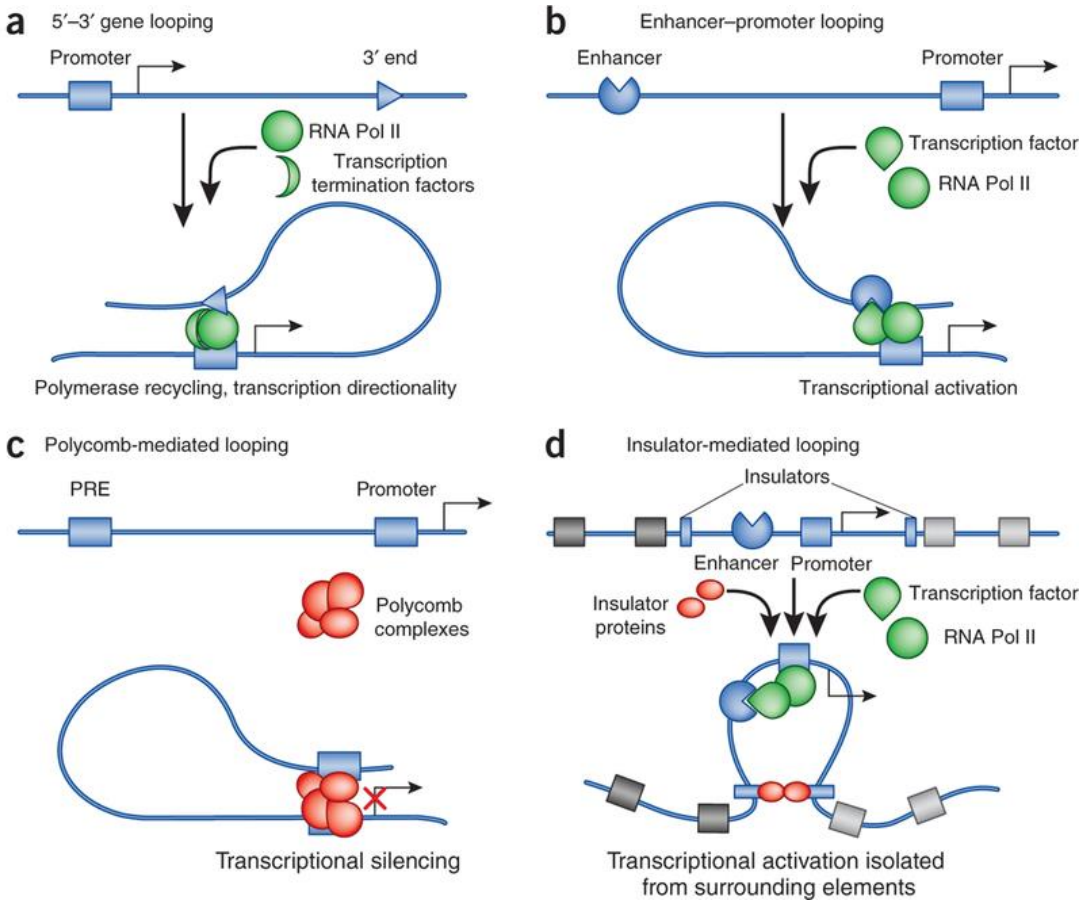


Figure 14: Chromatin loops and transcription regulation

(a) Intragenic loops promote RNApol recycling thus increasing transcription. (b) and (c) loops can bring a regulatory element (enhancer in (b) and silencer in (c)) in close proximity to its target promoter. (d) Loops can also serve as insulators limiting the diffusion of DNA supercoils (adapted from Cavalli et Misteli 2013).

While apparently absent from *Sulfolobus* crenarchaea (Takemata and Bell 2021), DNA loops have been recently observed in some archaea belonging to the Euryarchaeota superphylum (Cockram et al. 2021). *Haloferax volcanii* and *Halobacterium salinarum* exhibit in Hi-C contact maps the plaid pattern characteristic of DNA loops. However, *T. kodakarensis*, another euryarchaeon does not exhibit this pattern, highlighting the diversity of genome organizations in Archaea, even within the same superphylum.

4. Chromosome architects

Chromosome condensation and organization is dependent upon different classes of proteins that bend, wrap, link and cover DNA. They cooperate to form the chromatin, a highly packaged and organized state of the DNA that can take several different forms across the tree of life. The main proteins families involved in the formation of chromatin are described in this section. The list of most known nucleoid-associated proteins and their properties is found in **Table 2**.

Histones are surely the most famous chromosome packaging proteins. These highly basic proteins exhibit a positive charge at physiological pH that confers them a strong ability to bind DNA. They are sharing a general structure called the "histone-fold", made of three α-helices, two short and one long, connected by loops (Arents et al. 1991). Ubiquitous in Eukaryotes (Soo and Warnecke 2021), histones assemble in an octamer made of two copies of each core histones H2A, H2B, H3 and H4 (Luger et al. 1997). The octamer constrains roughly 150 bp DNA by wrapping it twice in a left-handed superhelical manner (**Figure 15**). Distribution of nucleosomes along the DNA molecule forms the famous beads-on-a-string structure (Woodcock, Safer, and Stanchfield 1976). The stability of the octamer is enhanced by the binding of the histone linker H1 at the entry and the exit of the nucleosome (Fyodorov et al. 2018). This H1-nucleosome structure is thought to contribute in Eukaryotes, to the further compaction of DNA in a 30 nm fibre (Thoma, Koller, and Klug 1979; Robinson and Rhodes 2006). N and C-terminal tails of core histones extrude from the octamer where they can participate to interactions between nucleosomes (Shogren-Knaak et al. 2006). These interactions are regulated by post-translational modifications (among others: acetylation, methylation, phosphorylation, ubiquitination, and biotinylation) that have a wide variety of effect on chromatin condensation (Tolsma and Hansen 2019). For instance, histone-tails acetylation inhibits nucleosome-nucleosome interaction, thus decondensing chromatin.

Table 2: Main characteristics of known nucleoid-associated proteins

Protein or group of proteins	DNA wrapping	DNA bridging	DNA bending	Binding motif	Molecular mass	Native protomer
Eukaryotes						
Core histones, H2A, H2B, H3 and H4	Yes	ND	ND	A ~10 bp periodic oscillation of AA/TT/TA elements in-phase with each other and out-of-phase with ~10 bp periodic GCs	11–14 kDa	Homodimer
Linker histones, H1 and H5	ND	Yes	ND	AT-rich DNA	~21 kDa	Homodimer
Smc	ND	Yes	ND	AT-rich DNA able to form secondary structures	~140 kDa	Heterodimer (for example, SMC1–SMC3)
Hmg	ND	ND	Yes	AT-tract sites	11–38 kDa	Homodimer or heterodimer (for example, HMG1–HMG2)
Euryarchaeota						
Archaeal histones HMfA and HMfB	Yes	ND	ND	$(A/T)_3NN(G/C)_3NN$	~7.5 kDa	Homodimer or heterodimer
Lrp	Yes	Yes	ND	ND	15–22 kDa	Homodimer
Alba	ND	Yes	ND	ND	~10 kDa	Homotetramer
MC1	ND	ND	Yes	AT-rich DNA	~10 kDa	Homodimer
HU	ND	ND	Yes	ND	~10 kDa	Homodimer
SMC	ND	Yes	ND	ND	~135 kDa	Homodimer
Crenarchaeota						
Lrp	Yes	Yes	ND	ND	~18 kDa	Homodimer
Cren7	ND	ND	Yes	ND	~7 kDa	Monomer
Sul7d	ND	ND	Yes	ND	~7 kDa	Monomer
Alba	ND	Yes	ND	ND	~10 kDa	Homodimer or homotetramer
SMC	ND	Yes	ND	ND	~100 kDa	Homodimer
CC1	ND	ND	ND	ND	~6 kDa	Monomer
Gram-negative bacteria						
HU	Yes	ND	Yes	A DNA structural motif in dsDNA joined to either dsDNA or ssDNA, with a mild preference for AT-rich or curved DNA	~9 kDa	Heterodimer (for example, HUα–HUβ)
Lrp	Yes	Yes	ND	(T/C)AG(A/T/C)A(A/T)ATT(A/T)T(A/T/G)CT(A/G)	~18 kDa	Homodimer
MukB	ND	Yes	ND	ND	~175 kDa	Homodimer
Fis	Yes (helically phased sites)	Yes	Yes	A_6 tracts and AT tracts	~11 kDa	Homodimer
H-NS	ND	Yes	ND	AT-rich DNA and TCGATAAATT	~15 kDa	Homodimer or heterodimer (H-NS–StpA)
IHF	ND	ND	Yes	(A/T)ATCAANNNNTT(A/G)	~11 kDa	Heterodimer (IHFα– IHFβ)
Dps	ND	ND	ND	ND	~19 kDa	Monomer or dodecamer
StpA	ND	Yes	ND	AT-rich DNA	~15 kDa	Homodimer or heterodimer (StpA–H-NS)
CbpA	ND	ND	ND	Curved DNA	~33 kDa	Homodimer or heterodimer (CbpA–CbpM)
CbpB	ND	ND	ND	Curved DNA	~33 kDa	Monomer
EbfC	ND	Suggested	ND	GTNAC	~11 kDa	Homodimer
MvaT	ND	Yes	ND	AT-rich DNA		Homodimer
Gram-positive bacteria						
MukB	ND	Yes	ND	Preference for ssDNA	~130 kDa	Homodimer
Lrp	Yes	Yes	ND	ND	~17 kDa	Homodimer
HU	ND	ND	Yes	ND	~10 kDa	Homodimer
Lsr2	ND	Yes	ND	AT-rich DNA	~12 kDa	Homodimer
Hlp	ND	ND	ND	ND	~21 kDa	Monomer
MrgA	ND	ND	ND	ND	~17 kDa	Monomer or dodecamer

Alba, acetylation lowers binding affinity; CbpA, curved-DNA-binding protein A; CbpB, curved-DNA-binding protein B (also known as Rob); CbpM, chaperone modulatory protein; Dps, DNA protection from starvation; dsDNA, double-stranded DNA; Fis, factor for inversion stimulation; Hlp, histone-like protein; Hmg, high mobility group; H-NS, histone-like nucleoid-structuring; IHF, integration host factor; Lrp, leucine-responsive regulatory protein; MrgA, metalloregulation DNA-binding stress protein; ND, not determined; Smc, structural maintenance of chromosome; ssDNA, single-stranded DNA.

This table lists characteristics from most of the known NAPs. Eucaryotic and archaeal histones and SMC are also included in the table (adapted from Dillon et Dorman 2010).

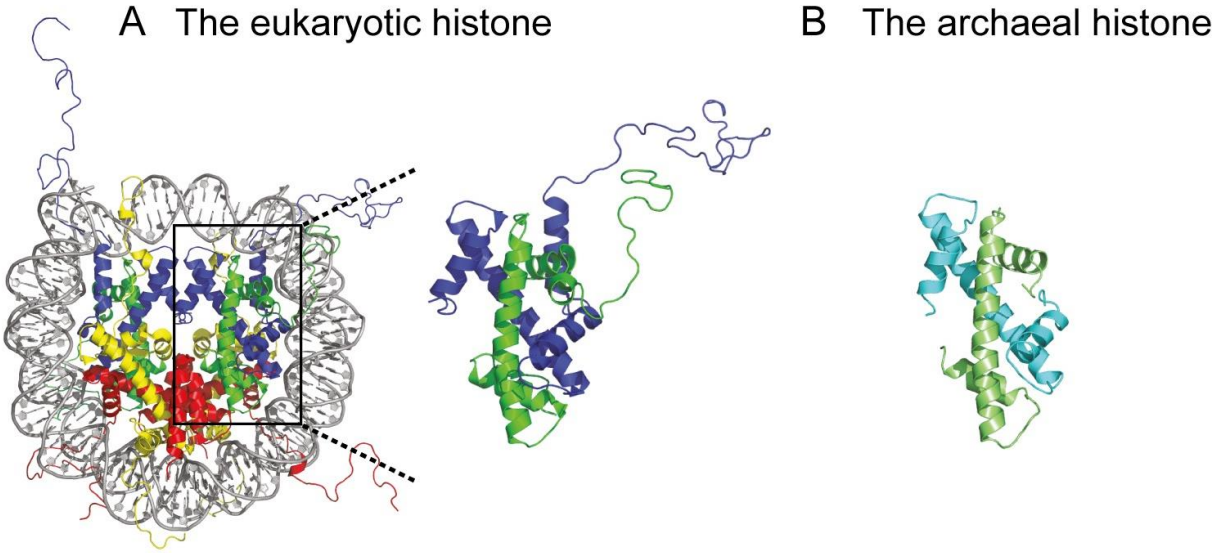


Figure 15: Archaeal histones are very similar to the eukaryotic ones

All known Eukaryotes encode histones. (A) Eukaryotic histones are organized into nucleosome, an octamer composed of two copies of H3 (blue), H4 (green), H2A (yellow) and H2B (red) histones that wraps DNA. (B) To the notable exception of Crenarcheota, Archaea also encode histones. Archaeal histones are homologous to H3 and H4 eukaryotic histones and share a very similar structure, but they lack the tails of eukaryotic histones meaning that they can not carry post-translational modifications (adapted from Henneman et al. 2018).

In contrast to Eukaryotes, Bacteria rely on a highly diverse class of DNA binding proteins, called nucleoid associated proteins (NAPs), to condense and organize their genomes (Dillon and Dorman 2010; Verma, Qian, and Adhya 2019; Hołówka and Zakrzewska-Czerwińska 2020). These small proteins (around 10-20 kDa) can bend, wrap, or bridge DNA molecules and constrain negative supercoiling. The most studied bacterial NAPs include the heat-unstable protein (HU), the integration host factor (IHF), the histone-like nucleoid structuring protein (H-NS), the factor for inversion stimulation (FIS), the Lrp (leucine-responsive regulatory protein) and the Dps (DNA-binding protein from starved cells). HU, IHF, Fis, and Dps constrain supercoiling by their ability to bend DNA, while H-NS and Lrp favour contacts between distal elements through bridging (H-NS) or wrapping DNA (Lrp) around multimers (Hołówka and Zakrzewska-Czerwińska 2020). The changes induced by NAPs on bacterial chromatin have deep

impacts on every DNA transaction, similarly to what is observed for histones in Eukaryotes. This is especially true for transcription that is directly sensitive to the NAPs-constrained supercoiling, and the ability of NAPs to control promoters and enhancers/silencers accessibility.

Archaea harbour more diverse chromatin organizations than Bacteria and Eukaryotes, based both on histones and NAPs. Indeed, histones are not solely present in Eukaryotes, almost every phylum of archaea except Crenarchaeota possess histones (Sandman et al. 1990; Henneman et al. 2018; Sanders, Marshall, and Santangelo 2019). Histones from archaea are homologous to H3 and H4 eukaryotic histones and share the same core of α-helices. But in contrast to eukaryotic histones, they are devoid of tails subjected to post-translational modifications (**Figure 15**). Moreover, early study of *Methanothermus fervidus* and *Methanobacterium thermoautotrophicus*, suggested that DNA is wrapped around histone tetramers that constrained only about 60 bp (Pereira et al. 1997). This was confirmed later on by micrococcal nuclease digestion coupled with deep sequencing (MNase-seq) in *H. volcanii* (Ammar et al. 2012) or by structural studies of *M. fervidus* histones with DNA (Mattiroli et al. 2017). Remarkably, archaeal nucleosomes can constrain either negative (left-handed) or positive (right-handed) supercoils similar to what is observed in Eukaryotes (Musgrave, Sandman, and Reeve 1991; Musgrave, Forterre, and Slesarev 2000). Another interesting feature of archaeal histones is their ability to organize in polymers of homodimers or heterodimers from various size beyond the initially described tetramer and instead of the classic eucaryotic octamer (Maruyama et al. 2013). This peculiar structure unique to archaea, can imply up to sixteen histone dimers in *T. kodakarensis*, covering 480 bp by 30 bp increment (Maruyama et al. 2013). Similar histone organization, now called "hypernucleosome" (**Figure 16**), has also been described for *M. fervidus* and predicted from amino-acid sequence similarities in other species suggesting that it is a general mechanism in Archaea (Henneman et al. 2018). This raised the possibility that

modulation of hypernucleosome length could be one of the main regulators of the transcription in Archaea, reminiscent to eukaryotic heterochromatin dynamics (Henneman et al. 2021).

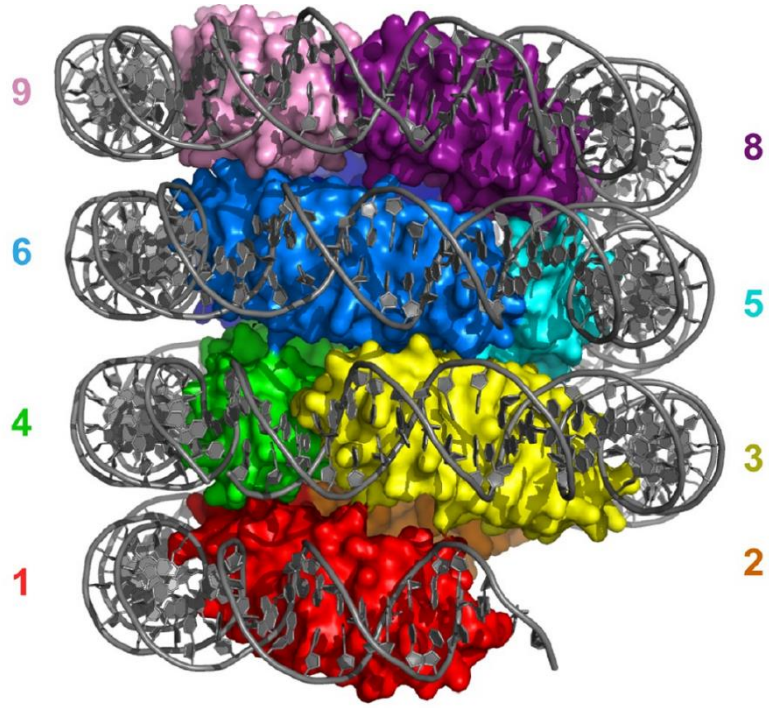


Figure 16: Structure of archaeal hypernucleosome

Nucleosome in archaea can be extended beyond the minimal histones tetramer by oligomerization of additional histones dimers, forming the hypernucleosome. In this structural simulation, nine HMfB histones dimers from *Methanothermus fervidus* are coloured (adapted from Henneman et al. 2018).

As in Bacteria, NAPs are numerous and highly diversified in Archaea (Zhang, Guo, and Huang 2012; Sanders, Marshall, and Santangelo 2019) (**Figure 17**). The most studied one is the acetylation lowers binding affinity protein (Alba) (Bell et al. 2002). This NAP was discovered in *Sulfolobus acidocaldarius* and then turned out to be widely spread in Archaea through the Sac10b family, especially in thermophiles (Green, Searcy, and DeLange 1983; P. Forterre, Confalonieri, and Knapp 1999). Alba can bind both DNA and RNA, thus playing a major role in genome organization and RNA stabilization. When bound to DNA, Alba can coat it and constrain negative supercoiling in addition to bridging distant sequences by its ability to dimerize. Potential regulation of Alba

function *in vivo* through post-translational modifications is still debated (Sanders, Marshall, and Santangelo 2019). Other highly studied archaeal NAPs include Sul7 and Cren7 specific for Crenarchaeota where they serve as substitutes for histones (Driessen and Dame 2011), as well as the HU homolog HTa in Thermoplasmata (Hocher et al. 2019). Some NAPs can also derive from transcription factors as the transcription regulator of the maltose-like system 2 (TrmBL2) (Lee et al. 2007; Wierer et al. 2016). This NAP is ubiquitous in Thermococcales and sporadically present in some DPANN and Bacteria. In Thermococcales, TrmBL2 is a major architect of the chromatin that competes with histones packaging by creating rigid protein-DNA filaments that cannot be transcribed (Maruyama et al. 2011; Efremov et al. 2015).

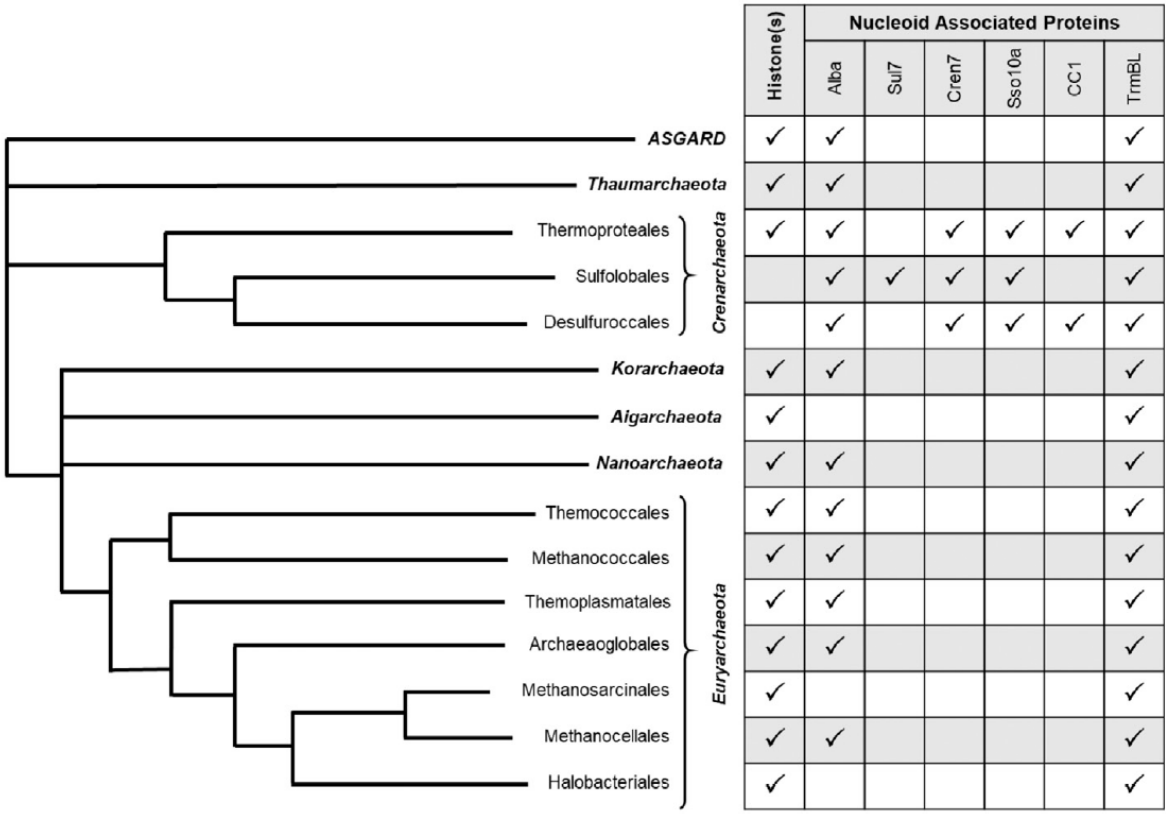


Figure 17: Distribution of NAPs and histones in Archaea

Histones are present in every archaeal lineage except Sulfolobales and Desulfurococcales. Alba and TrmBL are almost ubiquitous while others NAPs are restricted to Crenarchaeota. A schematic tree of archaea is indicated on the left (adapted from Sanders, Marshall, et Santangelo 2019).

A last major class of chromatin organizer is found in every living organism. Structural maintenance of chromosome complexes (SMC) have a ring shape made of two interacting ATP-binding sites and long coiled-coil structures, that stick through a hinge domain located at the top of the coiled-coil (Uhlmann 2016). There are two main classes of SMC, the cohesins and the condensins. As announced by their names, condensins are involved in genome condensation while cohesins maintain overall genome topology by forming topological barriers and keeping chromosomes together during replication. Recently, a new type of SMC called "coalescin" has been described in *Sulfolobus* (Takemata, Samson, and Bell 2019). Coalescin is responsible in this lineage for the formation of two transcriptional compartments, reminiscent of the A and B compartments in Eukaryotes (see section II. A. 1. for more details).

B. DNA supercoiling

1. Opening of the double helix creates topological constraints

For the propagation of organisms, genetic information must be transferred to new generation of cells. The discovery of DNA structure suggested to Watson and Crick a semi-conservative mechanism of DNA replication, that was only experimentally proved 5 years later (Meselson and Stahl 1958). This mechanism implies the opening of the double-helix to allow synthesis of new complementary strands according to canonical base-pairing possibilities. More generally, DNA transactions like replication, transcription or recombination all require the opening of the double helix. Because of their circular state and/or the contact with proteins, DNA ends are not free to rotate during these processes. For instance, the growing RNA attached to RNA-polymerase during transcription generates a drag force thus braking DNA rotation (Koster et al. 2010) (**Figure 18**).

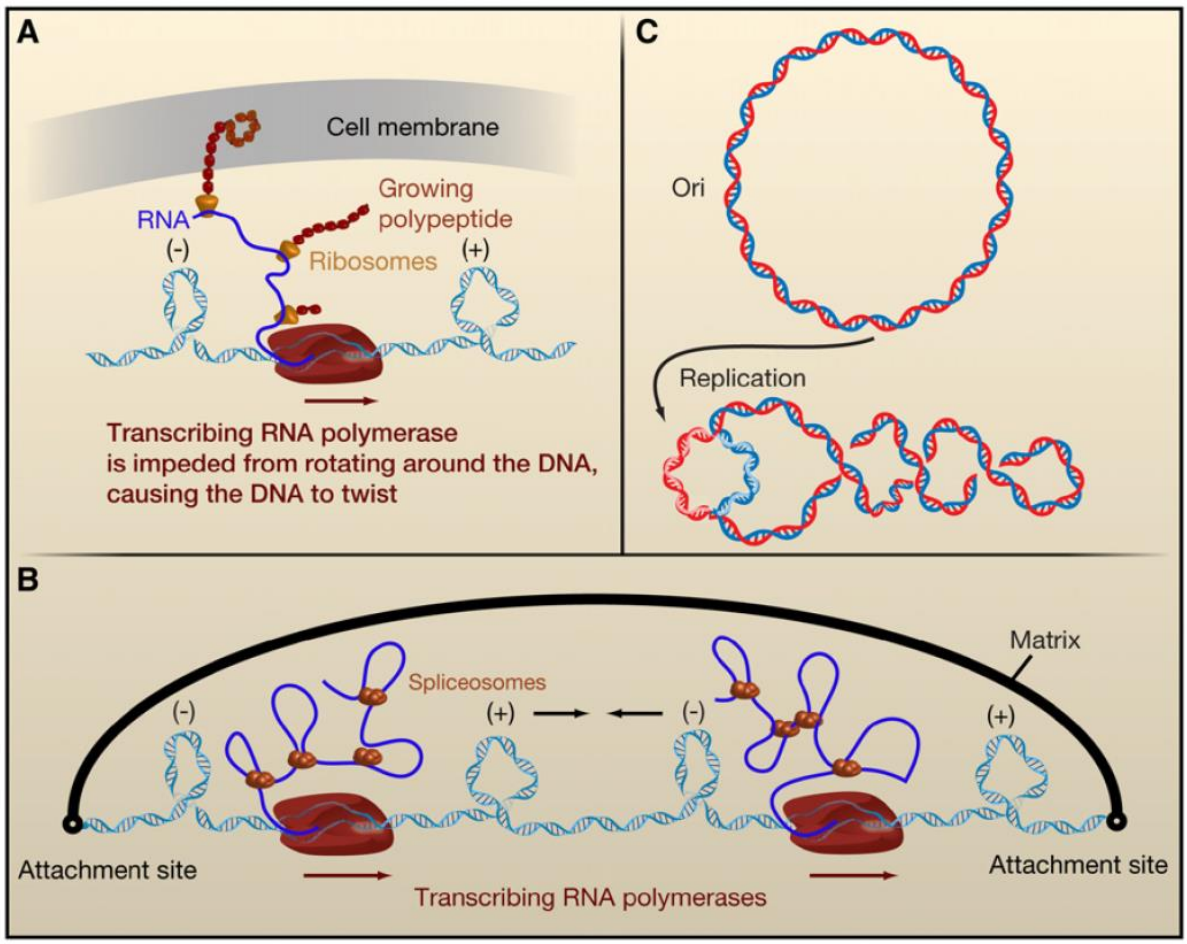


Figure 18: Topological consequences of transcription and replication

(A) In prokaryotes, RNA polymerase is not free to rotate during transcription because of nascent RNA and translating ribosomes that create a drag force. This results in the accumulation of positive supercoiling ahead of the polymerase and negative supercoiling behind. This phenomenon is even stronger when the newly synthesized peptide is anchored to the membrane. (B) Transcription and translation are not concomitant in Eukaryotes. But like in Bacteria the nascent RNA is creating a drag force. This drag force is enhanced by the bulky spliceosome and results in accumulation of negative supercoiling behind the RNA polymerase and positive supercoiling ahead. Supercoils from opposite senses can annihilate by diffusion during the transcription of tandem genes oriented in the same way. Conversely, supercoiling of identical sense generated in between divergent or convergent genes during transcription can accumulate. (C) Similarly, DNA helix opening and progression of the replication machinery cause positive DNA supercoiling accumulation ahead of the forks during replication. As the newly synthesized strand is not topologically closed negative supercoiling can not accumulate behind the replication machinery (adapted from Koster et al. 2010).

This phenomenon is even stronger in prokaryotes, where replication and transcription are coupled, creating a bulkier scaffold, or when a membrane protein is produced, anchoring the DNA to the membrane. A less obvious case is the possibility for nascent RNA to intertwine with DNA during transcription, creating a chimeric local duplex (R-loop by opposition to D-loop occurring with DNA) (Drolet 2006) that is taking the place of the complementary strand on DNA helix. The resulting constrained DNA is put under mechanical tension anytime the double helix is opened. These tensions could be lethal if they are not maintained to a level compatible with DNA transactions. Cells dispose of plenty of actors dedicated to this purpose, that are detailed further in this manuscript (section II. C. 1.).

2. Topological parameters

The formalization of DNA three-dimensional structure transitions and the action of cell actors to deal with them, gave rise to a new field, the DNA Topology. While complex in many respects, topology of a constrained DNA molecule relies on a relatively simple equation: Lk = Tw + Wr. In this equation, Lk is the number of topological links that can be defined as the number of times that one strand is crossing the other one for a DNA constrained in one plane; Tw is the number of twists that only depends on the size of the molecule (N) in bp and the pitch of the double helix (h), Tw = N/h ; Wr is the number of writhes that correspond to the number of times that the axis of the double helix is crossing the DNA plane. Practically, Tw reflects the number of coils and Wr the number of supercoils. For relaxed DNA molecule Wr is null. Overwinding introduces positive supercoils (overtwisting) while underwinding introduces negative supercoils (untwisting). These supercoils can be free or constrained by proteins. This leads to two kinds of supercoiling: (i) the toroidal where DNA is wrapped around a protein, typically

a histone ; (ii) the plectonemic where DNA is interwound, which is the classical state of circular supercoiled DNA in solution (**Figure 19**).

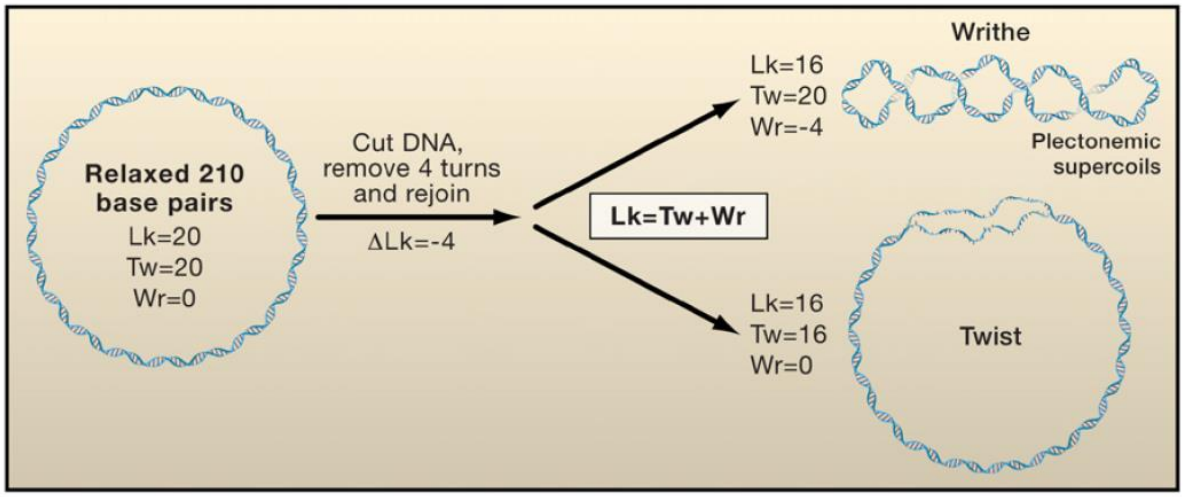


Figure 19: Description of DNA topology parameters

For a close circular DNA molecule or a linear one constrained at its extremities, DNA topology is described by the linking number (Lk) which corresponds to the number of time that one strand of the DNA crosses the other one in a planar projection. So Lk is the sum of twist (Tw) and writhe (Wr). For a topologically closed molecule Lk is a constant, meaning that every variations of Tw or Wr is automatically compensate by the other parameter. Tw depends on the length and the helix pitch of the DNA molecule. Assuming a helix pitch of 10.5 bp/turn, a 210 bp DNA, Tw = 20. Wr corresponds to the number of supercoils. For Wr = 0 the molecule is relaxed. Any variation of Lk implies introduction of a transient break into DNA (typically by a topoisomerase). In this example, 4 twists are removed from the relaxed molecule (ΔLk = -4). This could result in the introduction of 4 negative supercoils (Wr = Lk – Tw = 16 – 20 = -4) or in the denaturation of DNA at the expense of Tw (Tw = Lk – Wr = 16 – 0 = 16) (adapted from Koster et al. 2010).

To compare the topological state of different DNA molecules it is convenient to use a value that is not dependant of their respective size. The superhelical density (σ) is commonly used for this purpose. This value is defined by the difference of topological links between a molecule and its relaxed state (Lk_0), normalized by the Lk_0: $\sigma = (Lk - Lk_0) / Lk_0$. By doing the approximation that the Tw is constant in average for a given molecule, this equation results in a measure of its degree of supercoiling normalized

by its length. The σ is positive for a positively supercoiled DNA and negative for a negatively supercoiled one.

3. Physiological consequences of DNA supercoiling

The sense of supercoiling is of special importance for DNA stability. The minimum energetic state for a DNA molecule is the relaxed state. In other words, every supercoil in a DNA corresponds to an excess of energy that could promote the opening of the double helix for a negative supercoiling or prevent it for a positive supercoiling, respectively, by decreasing or increasing base stacking. Thus, supercoils are a precious reserve of energy for the cell favouring or preventing DNA transactions to occur.

This means that all proteins that require DNA melting are more likely to interact with negatively supercoiled DNA. But independently of DNA melting, DNA supercoiling can change distances and local conformations between genetic elements. This can be exemplified by the complex relation between supercoiling and transcription. Negative supercoiling acts as a global regulator of gene expression. In many ways, negative supercoiling can influence the transcription process. (i) The recruitment of RNAP on promoter *via* σ factors in Bacteria or transcription factor B (TFB/TFIIB) in Archaea and Eukaryotes, rely on their interaction with -10 and -35 elements. Supercoiling can modify both the distance between these elements and their orientation on the DNA helix, thus changing their potential of interaction with the RNAP-complex. (ii) Negative supercoiling promotes DNA melting, a required step for transcription initiation and DNA decoding. This is especially important for G-C rich promoters, as those involved in the stringent response in Bacteria. (iii) In Prokaryotes as well as in Eukaryotes, transcription can be dependent upon regulatory sequences called enhancer or silencers. Negative supercoiling acts by bringing together an enhancer or a silencer and its target promoter with consequences for gene expression regulation (Dorman 2006).

At high temperature, molecular motion is thought to replace negative supercoiling as major DNA melting factor (Forterre 2002). Thus, negative supercoiling is theoretically not essential for DNA transactions and could even be deleterious by promoting excessive DNA denaturation. On the other hand, positive supercoiling can become advantageous at high temperature by stabilizing double-stranded DNA. However, *in vitro* work produced contradictory results about specific DNA topology requirements for transcription at high temperature. Transcription at 78°C from *Sulfolobus* cell-free transcription system is not sensitive to the template DNA topology (Bell et al. 1998). Conversely, the cell-free transcription system from *Pyrococcus furiosus* is more efficient on negatively supercoiled DNA at 70°C and 90°C (Hethke et al. 1999). But no comparative analysis of promoters structures was performed to determine whether they can explain the observed differences instead of transcription machineries specificities. In addition, *in vivo* plasmid topology studies highlighted species-specific superhelical density variations in response to growth-phase and temperatures (López-García and Forterre, 1997, 1999) but underlying molecular mechanisms remain elusive.

Beyond its importance for transcription, supercoiling is also implied in the condensation of chromosome. While almost all the supercoiling is constrained by nucleosomes in Eukaryotes (toroidal supercoiling), in *E. coli* only about 60% of supercoiling are constrained and the 40% remaining are free (plectonemic supercoiling) (Bliska and Cozzarelli 1987). Plectonemic supercoiling is not compacting DNA efficiently (2.5 fold) compared to toroidal supercoiling around NAPs and nucleosomes and higher structures of chromatin, however both types of supercoiling (plectonemic and toroidal) should be considered for genome compaction (Boles, White, and Cozzarelli 1990).

Considering the importance of DNA supercoiling in genome stability, DNA transactions and genome compaction, its level must be regulated in cells. A class of enzymes called topoisomerases is dedicated to this purpose.

C. Regulation of DNA supercoiling

1. Topoisomerases

Topoisomerases are key enzymes that solve topological constraints in DNA. They all share the same kind of activity: (i) DNA cleavage by the formation of a transient phosphodiester bond between the DNA and a catalytic tyrosine of the topoisomerase ; (ii) passage of one or two DNA strands through the cleavage, or, in some cases, controlled rotation of the extremities produced by DNA cleavage; (iii) resealing of the cleaved part. This mechanism is changing the Lk of DNA molecules, allowing relaxation/supercoiling, knotting/unknotting, catenation/decatenation and duplex formation of DNA (**Figure 20**) (Bates and Maxwell 2005)

Topoisomerases families were named from I to VIII according to the order by which they were discovered. By chance, this order is coherent with their mechanistic activities that involved the cleavage of one strand for odd topoisomerases (Type I topoisomerases) or two strands for even topoisomerases (Type II topoisomerases). This principle has been conserved for the naming of all more recently discovered topoisomerases (**Table 3**).

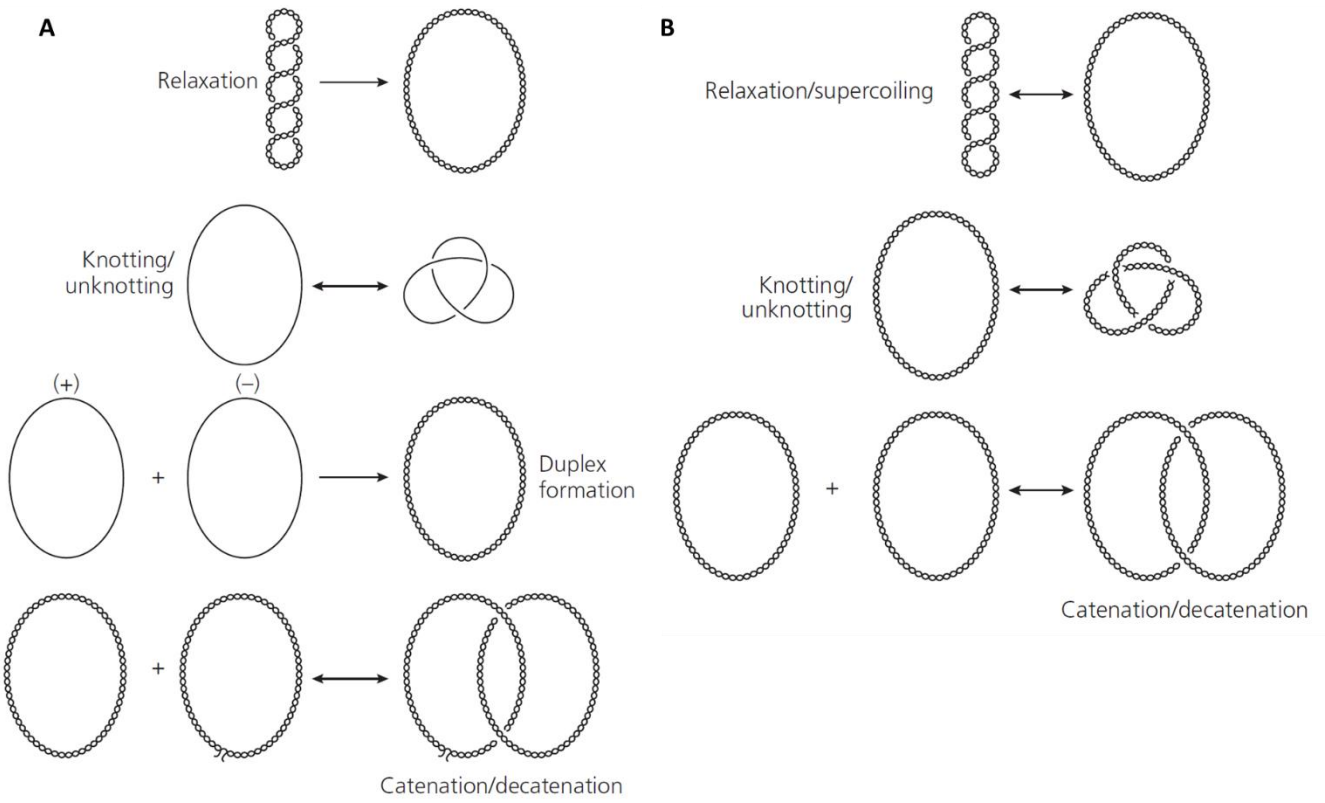


Figure 20: Reactions catalysed by topoisomerases

Schematic representation of (A) type I and (B) type II topoisomerases activities on topologically closed DNA. As type I toposiomerases cleave only one strand at a time, they need single-stranded or nicked DNA to perform the same reactions as type II topoisomerases.

Table 3: Main characteristics of known topoisomerases

Type	Family	Sub-family	Characteristics				Activities						Distribution		
			Enzyme structure	Bond formed	ATP-dependent	Mg2+-dependent	Relaxation		Supercoiling		Catenation/Decatenation	« Knotting/Unknotting »	Archaea	Bacteria	Eucaryotes
							Sur-	Sur+	Sur-	Sur+					
I	IA	I	Monomer	5’	No	Yes	+++	-	-	-	+	+	Yes	Yes	No
		III	Monomer	5’	No	Yes	+	-	-	-	++	+	Yes	Yes	Yes
		Reverse Gyrase	Monomer	5’	Yes	Yes	+++	-	-	+++	ND	-	Yes	Yes	No
	IB	IB	Monomer	3’	No	No	+	+	-	-	-	+	Yes	Yes	Yes
	IC	V	Monomer	3’	No	No	+	+	-	-	-	ND	Yes	No	No
II	IIA	II	Homo-dimer	5’	Yes	Yes	+	+	-	-	+	+	No	No	Yes
		IV	Hetero-tetramer	5’	Yes	Yes	+	+	-	-	+	+	No	Yes	No
		Gyrase	Hetero-tetramer	5’	Yes	Yes	-	+++	+++	-	+	+	Yes	Yes	No
	IIB	VI	Hetero-tetramer	5’	Yes	Yes	+	++	-	-	++	ND	Yes	No	ND
		VIII	Homo-dimer	ND	Yes / No	Yes	+	+	-	-	+	ND	Yes	Yes	No

Table 3: Main characteristics of known topoisomerases

(Sur-) DNA negatively supercoiled; (Sur+) DNA positively supercoiled; (Bond formed) covalent link to the 5’ or 3’ extremity of DNA; (ND) non-determined; (+) reaction catalysed (the number of + indicates the efficiency to catalyse the reaction); (-) non-catalysed reaction; (*) some topoisomerases VIII are ATP-dependent, others are not (adapted from Bush, Evans-Roberts, et Maxwell 2015).

2. Supercoiling homeostasis

Topoisomerases have many different functions in the cell. By their transient cleaving activity, they perform decatenation during replication, unknotting of entangled DNA and relaxation of supercoiled DNA. This last activity is of special importance for all DNA transactions. For example, transcription, replication or recombination can only occur if the supercoiling level is compatible with DNA helix opening.

In Bacteria, superhelical density is maintained around -0.06 in exponential growth phase, with about 40% (-0.025) of free supercoils. The remaining 60% (-0.035) are constrained, mainly by histone-like proteins (Bliska and Cozzarelli 1987). In Eukaryotes, DNA seems to be supercoiled in the same extent but with almost all the supercoils constrained by nucleosomes (Sinden, Carlson, and Pettijohn 1980). Unconstrained DNA is relaxed in average, with hot-spots of supercoiling near transcription start sites (TSS) of expressed genes (Kouzine et al. 2013; Krassovsky, Ghosh, and Meyer 2021).

The negative superhelical density value of -0.06 favours DNA melting and thereby allows DNA transactions. This implies that DNA supercoiling must be tightly regulated to fit within a narrow range compatible with different DNA-templated processes. These processes, however, by themselves directly modify DNA supercoiling. For instance, transcription requires DNA strands separation for DNA decoding by the RNA polymerase (RNAP) and ribonucleotides incorporation. The nascent mRNA causes molecular crowding/drag that prevents the free rotation of the RNAP around the DNA. Consequently, it is the DNA that is forced to rotate to compensate for the mechanical tension that is introduced by transcription. The result is the accumulation of positive supercoils ahead and of negative supercoiling behind of the RNAP, as proposed in the "twin supercoiled-domain" theory (Liu and Wang 1987). This phenomenon is enhanced by everything that can increase the drag force of the nascent mRNA, such as transcript splicing in Eukaryotes or coupling of translation and transcription in Prokaryotes. In the

latter case, translation of proteins bound to the membrane, creates a strong anchor that will completely block RNAP rotation (Koster et al., 2010).

DNA replication can be cited as a second example of a process in course of which, as the replication machinery extends the new-synthesized strand, DNA template must be opened. In a similar way to transcription, because the machinery is not free to rotate along DNA, positive supercoiling accumulates ahead of the replicative fork. But in contrast to transcription, negative supercoiling can not accumulate behind the replicative fork as the newly synthesized strand has a free extremity (Postow et al. 2001).

In both cases, the resulting accumulation of supercoiling acts as a physical barrier that can stop replication, transcription, and any other DNA transactions requiring DNA melting. Thus, it is critical for the cell to constantly remove excessive supercoiling. Some topoisomerases are dedicated to this purpose. In Bacteria, topoisomerase I (topo I, also called ω or swivelase) removes negative supercoiling while DNA gyrase (hereafter gyrase) introduces negative supercoiling. Their antagonistic activities are balanced by transcriptional regulation. The promoter of *topA*, encoding topoisomerase I, is induced by negative supercoiling and conversely, *gyrA* and *gyrB* promoters are induced by DNA relaxation. This leads to a homeostasis of supercoiling where the two main regulators, topoisomerase I and gyrase, through their activity directly up-regulate the expression of their antagonist (**Figure 21**) (Menzel and Gellert 1983).

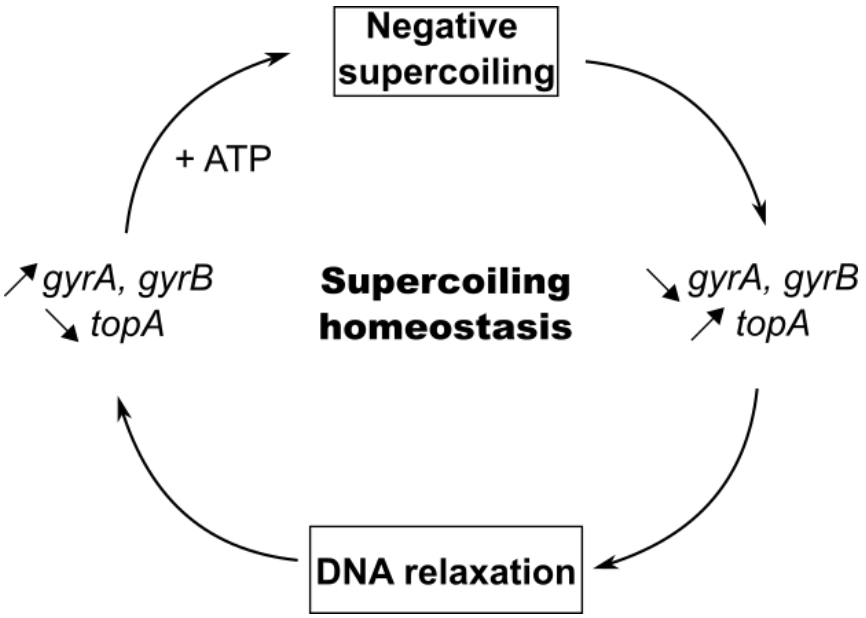


Figure 21: Model of supercoiling homeostasis in Bacteria

Negative supercoiling increases the transcription of *topA* by promoting the melting of its promoter, leading to DNA relaxation by topoisomerase I. Conversely, DNA relaxation increases the transcription of *gyrA* and *gyrB*, leading to DNA supercoiling by gyrase in presence of ATP.

Such a homeostatic regulation has not been yet described in Eukaryotes but has been recently proposed in archaeon *S. solfataricus*, based on *in vitro* experiments (Couturier et al. 2020) (**Figure 22**). From these experiments, topoisomerase VI (Topo VI) and the reverse gyrase TopR1 (*S. solfataricus* has two different reverse gyrases) are respectively assuming reminiscent roles of gyrase and Topo I in Bacteria. TopR1 removes excessive negative supercoiling by its activity of positive supercoiling while the Topo VI relaxes positive supercoiling. The other reverse gyrase TopR2 of *S. solfataricus* does not seem to be involved in this homeostatic mechanism, as well as the Topo III.

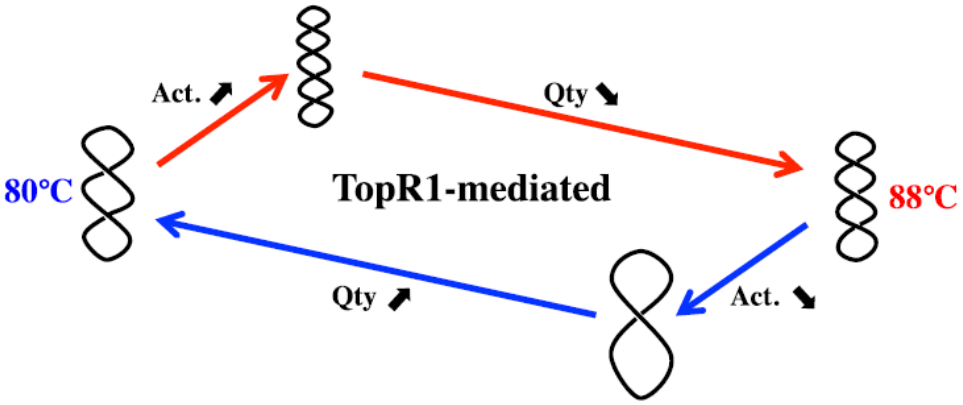


Figure 22: Model of supercoiling homeostasis in Sulfolobus

TopR1 is proposed as the main actor responsible of supercoiling homeostasis in *Sulfolobus*. Following temperature up-shift (red part of the schema) TopR1 activity increases, leading to accumulation of positive supercoiling and down-regulation of *topR1* expression. A temperature down-shift (blue part of the schema) results in the opposite behaviour: decreasing activity of TopR1 and up-regulation of *topR1* expression (adapted from Couturier et al. 2020).

III. Gyrase and reverse gyrase: two atypical topoisomerases

Among all topoisomerases, gyrase and reverse gyrase distinguish themselves by their unique activities of active supercoiling. They are not removing tensions by passive DNA relaxation, instead, they are introducing supercoils into DNA through ATP-dependant processes. This makes these enzymes of special interest to quickly change genome supercoiling level in response to environmental changes and cell's requirements.

A. DNA gyrase

1. An ubiquitous and critical enzyme in Bacteria

In 1976, Gellert and collaborators observed that phage λ can recombine with relaxed closed-circular substrates only in the presence of ATP and *Escherichia coli* cell extract. Knowing that this reaction can only occur on negatively supercoiled DNA, they presumed that something in the *E. coli* cell extract can introduce negative supercoils in these relaxed substrates. They purified the responsible enzyme, that they called DNA gyrase, and demonstrated that its activity is strictly ATP and $MgCl_2$ dependant (Gellert, Mizuuchi, et al. 1976). Following studies established that gyrase is a type IIA topoisomerase, capable of (i) relaxing positively supercoiled DNA by its negative supercoiling activity (Gellert, O'Dea, et al. 1976; Sugino et al. 1978) ; (ii) decatenating and unknotting entangled substrates (Mizuuchi et al. 1980; Liu, Liu, and Alberts 1980; Kreuzer and Cozzarelli 1980) ; (iii) relaxing negatively supercoiled substrates in absence of ATP (Gellert et al. 1977; N. P. Higgins et al. 1978; Sugino et al. 1978). Even though gyrase was shown to have versatile *in vitro* activities, its main *in vivo* activity is the relaxation of positively supercoiled substrate by introduction of negative supercoils. Indeed, in Bacteria, decatenation is mainly ensured by the topoisomerase IV, that is 100-fold more efficient than gyrase (Zechiedrich and Cozzarelli 1995). Moreover, gyrase is 10-fold more efficient in removing positive supercoiling than introducing it in relaxed substrate and has a higher affinity for positively supercoiled substrates (Ashley et al. 2017).

After its discovery in *E. coli*, it turned out that the gyrase is present in almost every bacterium, except the hyperthermophile *Aquifex aeolicus* that encodes a gyrase which lost the negative supercoiling activity (Tretter, Lerman, and Berger 2010). The ubiquitous distribution in bacteria and its essential function, makes gyrase a major antibiotic target (Bush, Evans-Roberts, and Maxwell 2015). Interestingly, gyrase is also present in some Eukaryotes, like plants and apicomplexan parasites that acquired this

enzyme from chloroplast and apicoplast, respectively, (Cho et al. 2004; Dar et al. 2007) and in Archaea where its distribution will be presented further in this manuscript (see section III.A.3).

Gyrase is composed of 2 GyrA subunits and 2 GyrB subunits, that form a A_2B_2 heterotetramer (Klevan and Wang 1980). GyrA (97 kDa) wraps DNA with its C-terminal part and cleaves DNA with its N-terminal part that contains the catalytic tyrosine. GyrB (90 kDa) binds ATP with its N-terminal part and interacts with GyrA with its C-terminal part. The C-terminal part of GyrB contains the TOPRIM domain responsible for the binding of Mg^{2+} ions that are necessary for the cleavage-religation reaction, while the C-terminal part of GyrA contains the GyrA-box motif (consensus sequence Q(R/K)RGG(R/K)G) (**Figure 23 and 24**). This motif is a defining signature of all gyrases and it causes a positive DNA bending that is required for gyrase supercoiling activity (Kramlinger and Hiasa 2006).

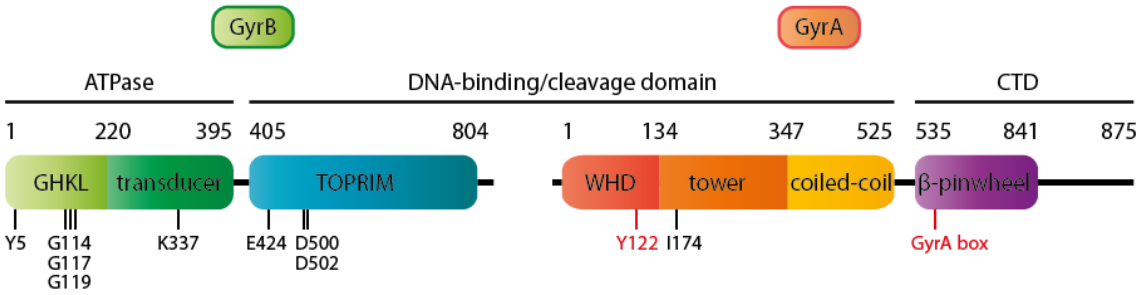


Figure 23: Primary domain structure of E. coli DNA gyrase

Schematic representation of the primary domain structure of GyrA and GyrB from *E. coli* including the functionally important residues. The GHKL domain hydrolyses ATP. The TOPRIM domain chelates three Mg2+ ions required for DNA cleavage and re-ligation reactions. The catalytic tyrosine (Y122) is in the WHD (winged-helix domain) domain. The C-terminal (CTD) part of GyrA contains the GyrA box which is essential for introducing the negative supercoiling in DNA by wrapping DNA around CTD domain.

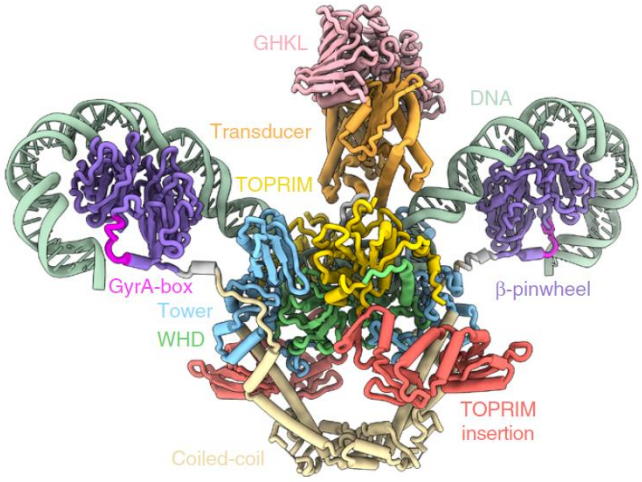


Figure 24: Molecular model of DNA Gyrase

E. coli gyrase interacting with 130 bp DNA was imaged by cryo-electron microscopy, then DNA and structural model of gyrase domains were built using the electron density map (adapted from Vanden Broeck et al. 2019).

The reaction of supercoiling starts with binding of the gyrase DNA-gate to DNA G-segment and wrapping the DNA around the C-terminal part of GyrA in a positive manner. Because of this wrapping, a part of the DNA molecule, called the T-fragment, locates just above the G-fragment. ATP binding to GyrB causes a conformational change trapping the T-fragment by closing the N-gate. The two catalytic tyrosines cleave the G-fragment by two nucleophilic attacks. Hydrolysis of one ATP molecule causes a GyrB rotation that transports the T-fragment through the cleaved G-fragment. Then, T-fragment is released at the C-gate, G-fragment is religated and finally, a second ATP molecule is hydrolysed to reset the gyrase (Hirsch and Klostermeier 2021) (**Figure 25**).

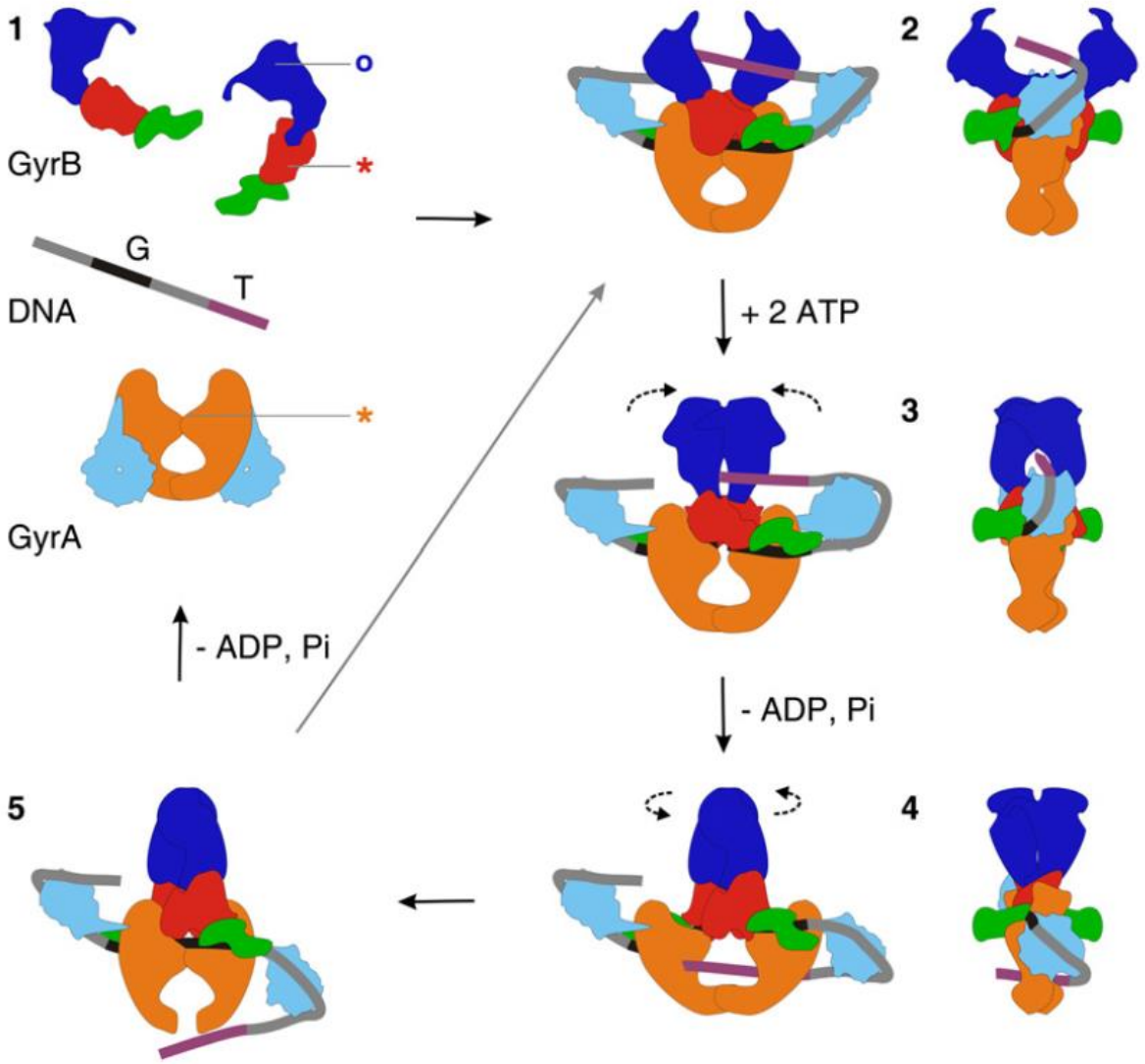


Figure 25: DNA gyrase supercoiling mechanism

GyrA CTD is light-blue and rest of GyrA is orange. GyrB tail is green, TOPRIM domain is red and the rest of GyrB is blue. DNA T segment is purple and G segment is black. Stars indicate the active site residues for DNA cleavage and the circle indicates the ATP-binding pocket. (1) Free state of gyrase subunits in solution. (2) Formation of the heterotetramer and wrapping of the DNA around GyrA CTD. T segment is positioned above G segment in a positive cross-over. (3) ATP-binding provokes GyrB clamp (N-gate) closing, capturing T segment while the G segment is transiently cleaved within DNA-gate. (4) Hydrolysis of one ATP allows GyrB to rotate, DNA-gate to open and the passage of the T segment through the cleaved G segment. (5) G-segment is religated, introducing two negative supercoils in the DNA molecule. Hydrolysis of a second ATP resets the enzyme before a new round of DNA supercoiling (adapted from Costenaro et al. 2007).

Gyrase is especially efficient in removing positive supercoils by its negative supercoiling activity. Because of that, gyrase plays a critical role in removing positive supercoils introduced by replication and transcription. In association with the topoisomerase I, it is responsible of supercoiling homeostasis (see section II.C.2).

2. Quick and massive gene regulation in response to environmental changes

DNA supercoiling is strongly sensitive to many environmental factors, including osmotic pressure, temperature, pH but also growth phase and aerobic versus anaerobic growth conditions(Higgins et al. 1988; Dorman et al. 1988; Dorman, Ni Bhriain, and Higgins 1990; Karem and Foster 1993; Cheung et al. 2003; Rui and Tse-Dinh 2003). Gyrase, through its ATP-dependent supercoiling activity, is directly involved in this topological response to environmental stress. Its activity is directly dependent of the ATP/ADP ratio and thereby of the cell's metabolic state (Drlica 1992). So, each variation of the environment that can modify the ATP/ADP ratio will change gyrase activity and consequently the expression of many genes sensitive to supercoiling (Dorman 2006).

This is a wide and quick way to adapt prokaryotic cells to changing environment. For instance, in *E. coli*, inhibition of gyrase and topoisomerase IV by drugs or temperature-sensitive mutant provokes a massive DNA relaxation. This deregulates up to 7% of *E. coli* genes (Peter et al. 2004). Interestingly, genes sensitive to supercoiling in *E. coli* are not evenly distributed in the chromosome, they are grouped in hot spots (**Figure 26**). Moreover, they can be as well up or downregulated, showing the true regulatory aspect of negative supercoiling beyond its basic propension to favour transcription. The same kind of study in *Streptococcus pneumoniae* also shows a massive supercoiling-dependent gene regulation (de la Campa et al. 2017). Chromosome relaxation by gyrase inhibition and accumulation of negative supercoiling by topoisomerase I inhibition, cause respectively the deregulation of 13% and 10% of *S. pneumoniae* genes, organized in hot spots as for *E. coli*.

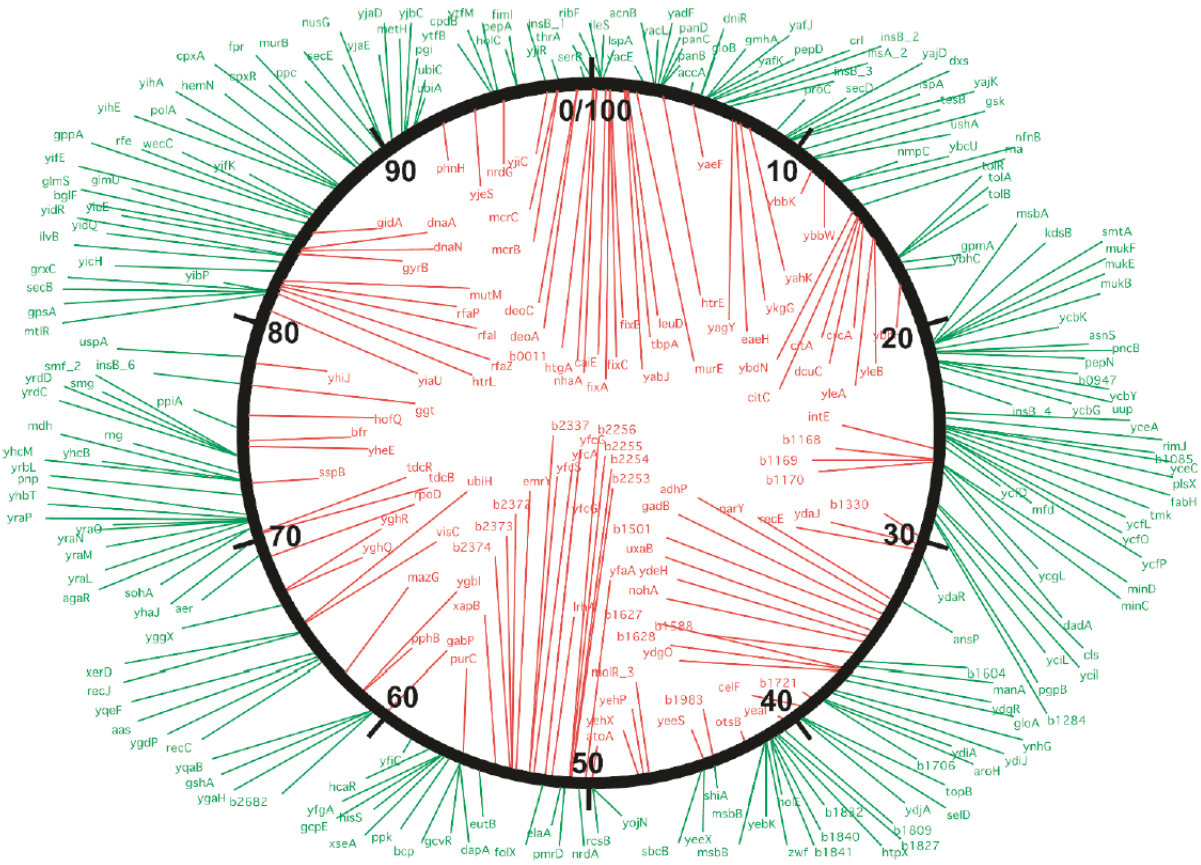


Figure 26: Chromosomal map of E. coli supercoiling sensitive genes

Loss of negative supercoiling consecutive to topoisomerase inhibition changes the expression of 7% of the genome (306 genes) located in hot spots all along the genome. Up-regulated genes are coloured in red and down-regulated are coloured in green (adapted from Peter et al. 2004).

Negative supercoiling could also be seen as a reserve of energy directly stored in the DNA (Cibot, n.d.). When the cell growth is stopped because of a nutrient lack or another kind of stress, the accumulation of negative supercoiling could be a powerful tool for the cell to restart its metabolism quickly. Negative supercoiling will lower the energy required by every DNA-melting process, boosting replication and transcription. The existence of such a mechanism for quick "emergency exit" is supported by the fact that in *E. coli*, *gyrA* and *gyrB* transcription increases in early stationary-phase and that the enzyme remains stable throughout long stationary-phase and is thus available to supercoil DNA as soon as ATP is produced again (Reyes-Domínguez et al. 2003).

Moreover, the survival of *E. coli* long stationary-phase cultures, that are maintained in appropriate conditions to preserve negative supercoiling (anaerobiosis and high NaCl concentrations), is enhanced by 5 log compared to normal culture after 5 days (Conter 2003).

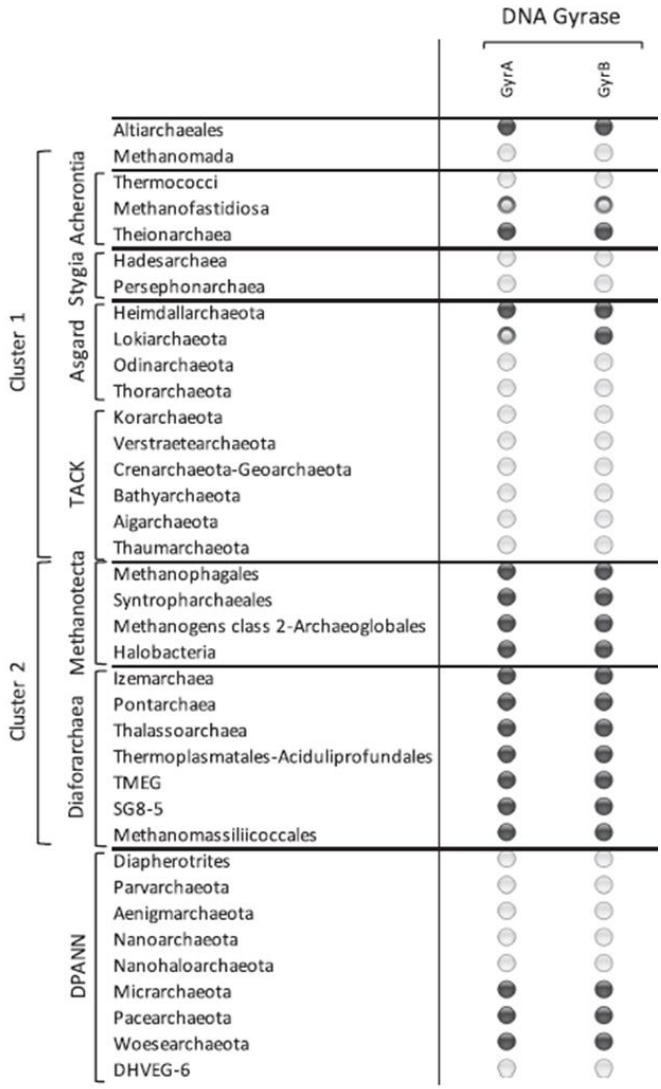
In summary, gyrase is ubiquitous in Bacteria, has essential role in overcoming positive supercoiling introduced by transcription and replication, and ability to regulate simultaneously the expression of many genes according to the metabolic state of the cell. By allowing Bacteria to quickly adapt to changing environment gyrase was proposed to be one of the key evolutionary inventions explaining the predominance of Bacteria in most biotopes (Forterre and Gadelle 2009).

3. Gyrase in Archaea

In Archaea, contrary to Bacteria, gyrase is not ubiquitous (**Table 4**). The first gyrase of archaea was discovered in Halobacteria and immediately used to set up selectable marker for genetic work, suggesting it is an essential enzyme in Archaea (Holmes and Dyall-Smith 1991). Then, gyrase was found in all members of group II Euryarchaeota, some DPANN and some of the recently discovered Asgards (Forterre et al. 2007; Raymann et al. 2014; Adam et al. 2017; Garnier et al. 2021). So far, the gyrase encoding genes were not detected in representatives of TACK superphylum or group I Euryarchaeota. Considering the importance of these groups in terms of number of lineages, a considerable part of Archaea does not carry gyrase. Phylogenetic analysis revealed that acquisition of gyrase by Archaea, likely occurred through one or perhaps two early HGTs from Bacteria, followed by gene losses in TACK and group I Euryarchaeota (Raymann et al. 2014). This acquisition surely had deep effects on DNA topology and cell physiology. Consequently, it was hypothesised that the benefit of

gyrase acquisition should have been strong and immediate to overcome the negative outcome of such a disturbing enzymatic activity (Forterre and Gadelle 2009).

Table 4: Distribution of gyrase in Archaea



Gyrase is present in all Cluster 2 Euryarchaeota and sporadically in some Asgard and DPANN archaea. Homologs of *gyrA* and *gyrB* were searched by Blast and HMM against representative genomes of archaea. Full circles represent gene presence in most or all members of the taxon, empty circles absence and partial circles presence in a few members only (adapted from Adam et al. 2017).

B. Reverse gyrase

1. An atypical topoisomerase

Among topoisomerases, gyrase is not the only one that can supercoil DNA. Another atypical topoisomerase, called by opposition reverse gyrase, has a positive supercoiling activity. But the comparison stops there. Reverse gyrase is not homologous to gyrase and is part of the type IA topoisomerases, together with the prokaryotic topoisomerase I and the topoisomerase III. Initially detected in cell extracts of *Sulfolobus acidocaldarius* (Kikuchi and Asai 1984), reverse gyrase results from the fusion of a SF2 helicase domain and a topoisomerase I domain (Confalonieri et al. 1993). The interaction of helicase and type IA topoisomerase is common in many organisms, including humans, with the topoisomerase III family and the helicase RecQ family frequently associated (Perugino et al. 2009; Chu and Hickson 2009). Such hybrid enzymes are involved in genome maintenance through replication, recombination and repair processes. Most reverse gyrases are in form of a single polypeptide but some of them have the helicase and the topoisomerase domains split in subunits that interact to form the active heterodimer (Kozyavkin et al. 1994; Capp et al. 2010; Catchpole and Forterre 2019). These split reverse gyrases are found in fast evolving species and are probably not the ancestral form of reverse gyrase (Lulchev and Klostermeier 2014).

N-terminal helicase part of the reverse gyrase binds and transiently unwinds DNA in an ATP-dependent manner (Ganguly, del Toro Duany, and Klostermeier 2013), making the reverse gyrase the only ATP-dependent type I topoisomerase. The C-terminal part of topoisomerase I-like domain is very similar to the topoisomerase I of *E. coli*. It consists of (i) a domain I that binds Mg^{2+} through the TOPRIM fold, (ii) a domain II that defines the central cavity of the enzyme, (iii) a domain III that contains the catalytic tyrosine and (iv) a domain IV that interacts with domains I and III to form the positively

charged interface for DNA binding, and to control the access to the central cavity by forming a gate that closed the central cavity (Schoeffler and Berger 2008) (**Figure 27**).

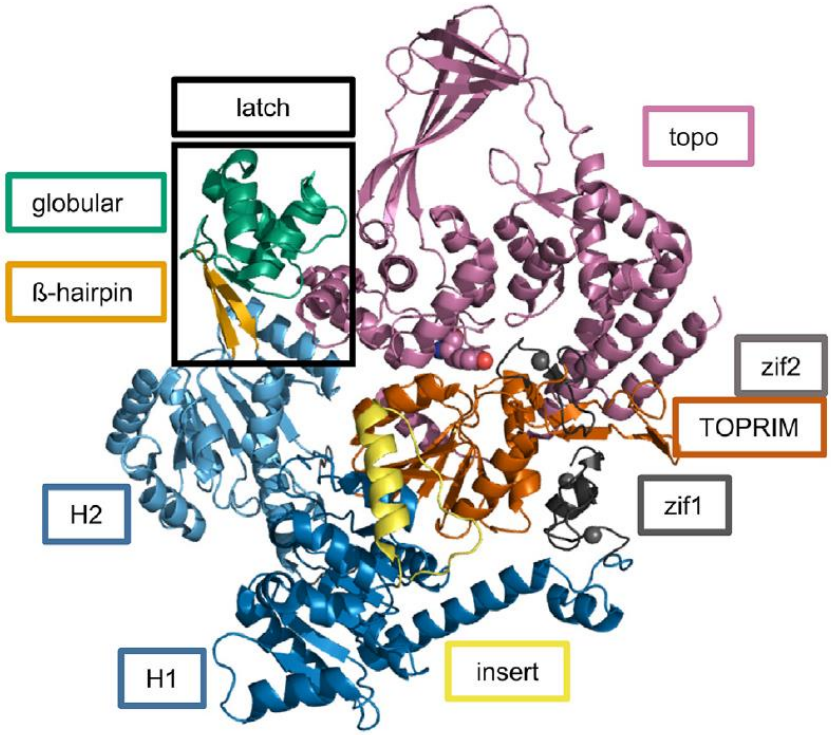


Figure 27: Structural model of Thermotoga maritima reverse gyrase

H1 and H2 (blue) form the helicase domain with an insertion (yellow) and the latch (green and orange). The TOPRIM domain (red) is indicated in the topoisomerase domain (pink) as well as two putative zinc fingers (zif1 and zif2) (adapted from Collin, Weisslocker-Schaetzel, et Klostermeier 2020).

Independently of the helicase domain, the topoisomerase domain can relax supercoiled DNA in ATP-independent manner (Perugino et al. 2009). The two independent helicase and topoisomerase activities led to the early model for reverse gyrase function. The model proposed that the DNA opening by the helicase activity alone creates negative supercoils behind the reverse gyrase and positive supercoils ahead while the enzyme moves along the DNA. Then topo I activity would assume the specific relaxation of negative supercoils, thus letting only the positive supercoils remain (Confalonieri et al. 1993). However, further studies proved that both domains

need to cooperate all along the process to supercoil DNA (Déclais et al. 2000). The cooperation between the helicase and the topoisomerase domains is dependent of a small insertion, called the latch (Rodriguez 2002). The latch is proposed to assure the communication between the two domains by opening the topoisomerase domain in response to DNA binding and ATP hydrolysis. These findings led to a new model where reverse gyrase segregates positive and negative supercoiling and relaxes specifically the negative one. In a third model, alternation of high affinity for ssDNA and dsDNA by the helicase domain, coupled with strand passage mechanism would result in positive supercoiling (del Toro Duany et al. 2008). Recently, the sequence of DNA topology transitions during the catalytic cycle has been uncovered by single molecule analysis (Yang et al. 2020) Despite these results, the exact catalytic cycle of reverse gyrase remains elusive.

2. The only hallmark of hyperthermophily

Life at high temperature is not possible without specific adaptations. Molecular motion increases proportionally to the temperature, making membrane, proteins and nucleic acids unstable (López-García et al. 2015). To overcome these difficulties hyperthermophiles have specific lipids (Koga 2012; Siliakus, van der Oost, and Kengen 2017), GC-rich rRNAs (Boussau et al. 2008), highly stable proteins (Zeldovich, Berezovsky, and Shakhnovich 2007), specific chaperones (Makarova, Wolf, and Koonin 2003) and specific ion transporters (Albers et al. 2001).

In the particular case of DNA, high temperature results in local denaturation that exposes nitrogenous bases to the solvent, increasing dramatically DNA damage such as depurinations, deaminations, oxidations, alkylations, and ultimately, double strand breaks (Lindahl 1993; Marguet and Forterre 1994). To face all these lesions, genomes of hyperthermophiles can rely on efficient DNA repair mechanisms (Grogan 2000; Ishino and Narumi 2015), DNA binding proteins (Sanders, Marshall, and Santangelo

2019), high amount of ions and polyamines (Grogan 1998; Morimoto et al. 2010; Terui et al. 2005) and the reverse gyrase (Forterre 2002).

From all the known actors involved in adaptation to thermophily, reverse gyrase is the only one shared by all hyperthermophiles and in most thermophiles but completely absent from mesophiles (Forterre 2002; Catchpole and Forterre 2019) (**Figure 28**).

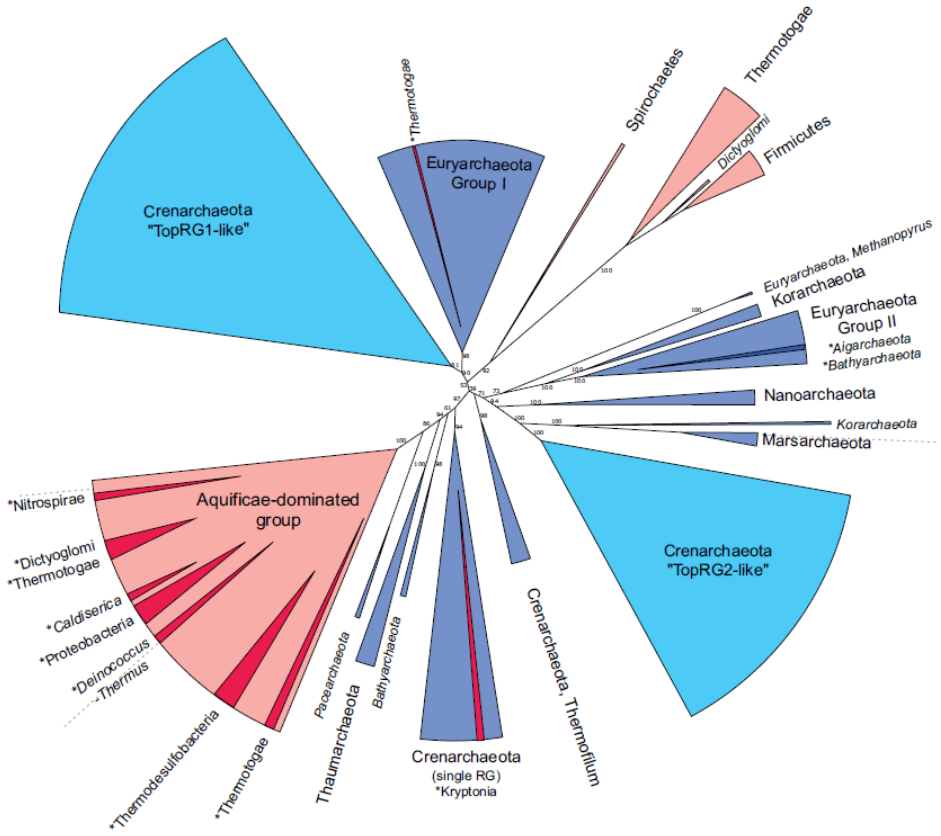


Figure 28: Phylogenetic tree of bacterial and archaeal reverse gyrase sequences

Archaeal clades are colored in blue, Bacterial clades in red, with phyla names indicated. Clades formed inside canonical phyla are indicated in darker shades, and labeled with an asterisk. Clades labelled with italicized text indicate that less than two sequences are present. Crenarchaeal TopRG1-like and TopRG2-like paralogues are indicated in pale blue. Ultrafast bootstrap values for major bipartitions are indicated on branches (adapted from Catchpole et Forterre 2019).

Thus, reverse gyrase is considered as the only hallmark of hyperthermophily. This strong association between thermophily and reverse gyrase can be reasonably explained by its activity of positive supercoiling, as overwound DNA is less sensitive to denaturation. But this simple explanation is challenged by a growing number of observations, leading to the proposition of other thermal protection mechanisms for reverse gyrase.

3. Uncertain *in vivo* activities

The different studies on the importance of reverse gyrase for life at high temperature have produced contradictory results. Reverse gyrase deletion is viable at optimal growth temperature (85°C) but lethal at 93°C for the hyperthermophilic archeon *T. kodakarensis* (Atomi, Matsumi, and Imanaka 2004). It is lethal for *Pyrococcus furiosus* at 95°C, a suboptimal growth temperature, and for *Sulfolobus islandicus* at 76°C its optimal growth temperature (Lipscomb et al. 2017; Zhang et al. 2018). This suggests that reverse gyrase is essential for hyperthermophilic lifestyle, thus corroborating its phylogenetic distribution. Interestingly, though, the DNA of hyperthermophiles is not always relaxed or positively supercoiled. While most of them, like Thermococcales and Sulfolobales do follow this classical rule (López-García and Forterre 1999; Charbonnier and Forterre 1994), other ones like Thermotogae bacteria and Archaeoglobales archaea have negatively supercoiled plasmid DNA (Guipaud et al. 1997; Bouthier de la Tour et al. 1998; López-García et al. 2000). The particularity of these organisms is to possess gyrase in addition to reverse gyrase. This means that gyrase activity is the dominant one when both enzymes coexist in the cell. Moreover, this suggests that underwound DNA is compatible with hyperthermophilic lifestyle and that the positive supercoiling activity of the reverse gyrase is not critical for hyperthermophiles.

This hypothesis is corroborated by several experimental observations. For instance, reverse gyrase of *Archaeoglobus fulgidus* prevents DNA strand breakage at apurinic

sites and keeps together broken extremities independently of its topoisomerase activity. In combination with the fact that reverse gyrase selectively binds nicked DNA, this led authors to propose that reverse gyrase acts *in vitro* as a DNA chaperone that stabilises DNA lesions until they are repaired (Kampmann and Stock 2004). Besides, reverse gyrase was shown to massively bind DNA *in vivo* after UV irradiation (Napoli et al. 2004), interact with single-strand binding protein (SSB), a common actor of DNA repair processes (Napoli et al. 2005) and with the translational polymerase Y (Valenti et al. 2009). Reverse gyrase also has a protective effect *in vivo* against double-strand breaks after methyl methanesulfonate treatment (MMS), an alkylating agent (Han, Feng, and She 2017). Taking together, these results strongly support a role of reverse gyrase in DNA repair, independently of its topoisomerase activity. Which reverse gyrase activity is really determinant for hyperthermophilic lifestyle is still an open question. Answering this important question would require decoupling DNA repair and DNA topology activities *in vivo*.

IV. Investigating DNA topology in Archaea: the special case of Gyrase

The acquisition of gyrase by an ancestral archaeon had, in principle, completely disturbed its original DNA topology, causing deep impacts on its genome architecture, DNA replication and gene expression. Nevertheless, the gyrase was established in many archaeal lineages suggesting that its introduction was somehow beneficial. One, and so far the only, attractive hypothesis is that this particular event could have participated to the adaptation of Archaea to mesophilic environment. In this chapter, I will first describe this theory and then present the experimental strategy I have set up to test this hypothesis.

1. Gyrase evolution in Archaea

As detailed above, DNA supercoiling is inherent to the very structure of DNA making topoisomerases crucial for life. This is emphasised by the presence of at least one type I and one type II topoisomerase in every living organism (Schoeffler and Berger 2008). Even though its accumulation is lethal, DNA supercoiling turned out to be advantageous in many situations through the high potential of genome condensation and gene expression regulation that it provides. Thus, topoisomerases are general regulators of DNA transactions by reducing the burden of excessive DNA constraints and enhancing the useful ones. Despite their importance, the evolutionary history of topoisomerases is still elusive. Many different classes of topoisomerases co-exist across species and their precise activities in the cells are often not defined. The complexity of the picture is even exacerbated by the paucity of data about global genome architecture and the importance of DNA supercoiling in most organisms.

Concerning the evolution of topoisomerases, some inter-domain transfers have been proposed, like acquisition of gyrase by Archaea or reverse gyrase by Bacteria (Raymann et al. 2014; Catchpole and Forterre 2019). Given the importance of supercoiling for genome structure and gene expression regulation, it is likely that the acquisition of a topoisomerase would have an important and global impact on the physiology of the concerned organism. This is especially true for the gyrase that dictates the overall genome topology in Bacteria by its ATP-dependent negative supercoiling activity (Forterre et al. 2007; Forterre and Gadelle 2009; Raymann et al. 2014).

Previously published phylogenetic trees suggest that gyrase was transferred once or twice from Bacteria to an ancestor of one group of Euryarchaeota and then lost in several lineages depending on the root for the archaeal tree (Forterre et al. 2007; Raymann et al. 2014). Other work, using reconstruction of ancestral sequences, suggested that the ancestral archaeon recipient of gyrase was a thermophile encoding a reverse gyrase, meaning therefore that its genome was probably relaxed or positively

supercoiled (Boussau et al. 2008; Groussin and Gouy 2011; López-García et al. 2015; Catchpole and Forterre 2019). In a "normal" mesophilic context, even modest changes in DNA topology would be deleterious for the cell, as demonstrated in *Streptococcus pneumoniae* and *E. coli* (Ferrándiz et al. 2016; Pruss, Manes, and Drlica 1982), but in the context of a thermophile the change from positive to negative supercoiling would correspond to drastic alteration of DNA topology with potentially lethal consequences. Indeed, in thermophiles positively supercoiled DNA is assumed to be important for protecting DNA from excessive melting and damage at high temperature. Thus, the acquisition of a gyrase by a naïve thermophile would potentially have dramatic consequences for genome stability. To overcome such an accumulation of negative effects, the gyrase recipient ancestor may have had disposed of strong resilience mechanisms against topological stress to maintain its DNA topology and/or had benefited of significant and immediate advantage procured by the gyrase activity (Forterre and Gadelle 2009).

2. Gyrase and adaptation to mesophilic lifestyle

More than 20 years ago Purification López-García emitted the hypothesis that the gyrase acquisition by a euryachaeal lineage was one of the key evolutionary events allowing these organisms to adapt to mesophilic conditions and thus conquer new ecological niches (López-García 1999; López-García et al. 2015). Indeed, published phylogenetic reconstructions all agree that the Last Archaeal Common Ancestor (LACA) was probably a thermophile (López-García et al. 2015; Catchpole and Forterre 2019). The transition from a thermophilic lifestyle to a mesophilic one, implies that the organisms need to compensate (because of temperature drop) for lower membrane fluidity/permeability, lower protein flexibility/activity and more difficult melting of DNA. It seems that this was done through the acquisition of numerous bacterial genes, involved in aerobic respiration (electron transport chains and cytochromes), protein folding (heat-shock protein Hsp70), membrane fluidity (fatty acid biosynthesis enzymes), membrane permeability (transporters) and DNA unwinding (gyrase) (López-

García et al. 2015). Gyrase, through its negative supercoiling activity, would be the key actor allowing DNA transactions to occur by lowering the energy required to open DNA helix. But gyrase is not the only topological actor that can provide negative supercoiling. The same goal can be reached with histones constraining negative supercoiling by wrapping DNA around the nucleosomes as it happens in Eukaryotes and Archaea (Musgrave, Forterre, and Slesarev 2000).

Whereas gyrase is present in almost every mesophilic archaeon, only few thermophilic lineages are equipped with a gyrase. This raises the question of an incompatibility, between the negative supercoiling generated by gyrase and life at high temperature. Still, thermophiles such as Archaeoglobi archaea (López-García et al. 2000) or Thermotogae bacteria challenge this hypothesis (Guipaud et al. 1997). In these organisms endogenous plasmids are negatively supercoiled, suggesting that these organisms have negatively supercoiled genomes.

V. Objectives of the thesis and experimental strategies

1. Reconstruct the evolutionary history of gyrase in Archaea

As we have seen above, the acquisition of gyrase by Archaea was perhaps one of the key events at the origin of their expansion in mesophilic environments. The discovery of new gyrase-encoding lineages outside of Euryarchaeota, like Asgards or DPANN, challenge the current view of a unique transfer of gyrase from Bacteria to Euryarchaeota. Indeed, these lineages are distantly related to Euryarchaeota and therefore the evolutionary scenario explaining the gyrase distribution becomes more complex, possibly involving multiple inter-domain and/or intra-archaea horizontal gene transfers. The first objective of my thesis was to perform a global and updated phylogenetic analysis of currently available bacterial and archaeal gyrase sequences.

Such analysis should allow us to determine more precisely the timing and number of horizontal gene transfers throughout the evolutionary history of archaea. This analysis coupled with investigation of associated genomic adaptations could provide a general overview of the required conditions for establishment of gyrase in a naïve organism. This, still ongoing, work is presented as draft of article I of my thesis entitled "Evolutionary history of archaeal DNA gyrases".

2. Describe the consequences of gyrase introduction in a naïve hyperthermophilic archaeon

Gyrase most certainly altered the original DNA topology of the ancestral thermophilic recipient archaeon. But what are the concrete consequences of introducing such an "opposite twist" in a naïve organism is not known. Is it really toxic? And if so why? Does the successful implantation of gyrase imply specific adaptations? Could it be sufficient to lower the growth temperature of this organism to compensate for gyrase activity? Is gene expression affected, and if so, what are the affected functions? To address these questions, I decided to mimic the ancestral transfer of a bacterial gyrase in a hyperthermophilic archaeon. Although, this strategy using a modern organism is not fully extendable to the ancestral situation it still can give precious insights about the basic molecular mechanisms which allow the cells to resist such a challenge.

I chose *Thermococcus kodakarensis* as the recipient organism for the gyrase genes. This marine, hyperthermophilic archaeon belongs to the Euryarchaeota superphylum. It lives in hydrothermal vents where it feeds on organic matter such as polysaccharides and peptides and it performes anaerobic sulphur-based respiration to produce energy. Similar to other Thermococcales, *T. kodakarensis* exhibits a single small circular chromosome of 2.08 Mb with 52% GC-content (Atomi et al. 2004). This organism became progressively one of the most studied model archaea because of its (i) natural competence, (ii) short generation time of about 40 min in rich medium at an optimal

growth temperature of 85°C, and because of (iii) the availability of a panel of genetic tools and replicative plasmids (Hileman and Santangelo 2012). Importantly, *T. kodakarensis* encodes two histone paralogs (HTkA and HTkB), two NAPs (Alba and TrmBL2), one SMC and three topoisomerases (reverse gyrase, topoisomerase III and topoisomerase VI). This topological kit should result in a slightly positive DNA superhelical density, comparable to other Thermococcales (~0.030) (Charbonnier and Forterre 1994; López-García and Forterre 1997), and a chromatin extensively covered by histones and NAPs.

I chose *Thermotoga maritima*, a hyperthermophilic bacterium as a donor for the gyrase genes. This bacterium occupies the same environments as *T. kodakarensis* (Huber et al. 2004). Importantly, plasmids from *T. maritima* are negatively supercoiled *in vivo* proving that the gyrase activity dominates over reverse gyrase (Guipaud et al. 1997). Moreover, *T. kodakarensis* and *T. maritima* can be co-cultured at 85°C which is optimal growth temperature for *T. kodakarensis*. I therefore hypothesised that the gyrase of *T. maritima* could be active in *T. kodakarensis*.

To test this hypothesis, I investigated gyrase activity in *T. kodakarensis* by superhelical density measurements on plasmids, through 2D agarose electrophoresis. To understand global impact of gyrase on *T. kodakarensis* cells I investigated how toxic was gyrase by (i) monitoring growth; (ii) measuring plasmid loss; (iii) measuring cell shape and size parameters using light microscopy. Finally, using RNA-seq I investigated how the gyrase expression and in particular its negative supercoiling activity affects the gene expression. The results of this work are reported in article II of my thesis entitled "The hyperthermophilic archaeon *Thermococcus kodakarensis* is resistant to pervasive negative supercoiling activity of DNA gyrase".

3. Investigate the importance of DNA supercoiling for adaptation to hyperthermophilic lifestyle in archaea

Since the discovery of DNA structure, the importance of DNA supercoiling for DNA transactions has been extensively studied in model bacteria and eukaryotes. It came out from these studies that a minimum level of negative supercoiling is required in these organisms to sustain transcription and replication. Bacteria reach such state mainly through a topoisomerase-based homeostatic regulation while Eukaryotes rely on histones to constrain DNA and provide toroidal negative supercoiling. Comparatively, very little is known in Archaea about the importance of supercoiling for DNA transactions. The results of few studies reported so far (see chapter II.B.3 Physiological consequences of DNA supercoiling) give a partial view of the importance of supercoiling in archaea in general and in hyperthermophiles in particular. This is an important caveat for understanding the basic mechanisms governing genome stability and functionality, and the specific contribution of reverse gyrase for life at high temperature.

Slow electrophoresis for separation of plasmid topoisomers has been used for decades to investigate the topological state of DNA *in vivo*. While useful in the frame of my thesis work to confirm gyrase activity in *T. kodakarensis*, this method can not give more than a global *a priori* for the chromosome topology. Indeed, by its size, its chromatinization and the complexity of transactions it undergoes, chromosome harbours a multitude of topologies distinct of the plasmid one. To go beyond this limitation and truly map the supercoiling in the chromosome of *T. kodakarensis*, I decided to use a method that relies on preferential crosslinking of trimethyl-psoralen in negatively supercoiled DNA coupled to next-generation sequencing (Kouzine et al. 2013; Lal et al. 2016; Kouzine et al. 2018). This technique alone is not sufficient to meaningfully interpret the mapping data. Indeed, histone occupancy also has to be taken into account and I have therefore used a well-established method called micrococcal nuclease digestion in combination with deep sequencing (MNase-seq).

This method allows precise mapping of nucleosomes across the genome by sequencing DNA protected from MNase digestion by nucleosomes (Schones et al. 2008). In the article III of my thesis manuscript, I described the final protocol for psoralen-crosslink in *Thermococcus* up to producing the sequencing samples. The oncoming analysis of sequencing data should allow me to finalise this part of my work. I also included in this analysis *T. kodakarensis* deleted for reverse gyrase. This should allow me to describe whether this topoisomerase functions in controlling the DNA topology at the global scale in *T. kodakarensis* and in hyperthermophiles more generally. In article IV of my thesis I reported the ongoing work on investigation of the DNA topology in this mutant.

Article I

Evolutionary history of archaeal DNA gyrases

Paul Villain[1], Ryan Catchpole[1,2], Patrick Forterre[1,3], Jacques Oberto[1], Violette da Cunha[1, $] & Tamara Basta,[1,$]

1. Université Paris-Saclay, CEA, CNRS, Institute for Integrative Biology of the Cell (I2BC), 91198, Gif-sur-Yvette, France
2. Present address : Department of Biochemistry and Molecular Biology, University of Georgia, Athens, GA 30602, USA
3. Archaeal Virology Unit, Institut Pasteur, Paris, France.

[$]Corresponding authors:
E-mail: tamara.basta@i2bc.paris-saclay.fr
E-mail: violette.da-cunha@i2bc.paris-saclay.fr

ABSTRACT

Replication, transcription and other DNA-templated processes that require opening of the double helix lead to topological constraints which if not resolved are lethal for cells. Topoisomerases are ubiquitous enzymes that solve these topological problems by introducing one (type I) or two (type II) transient strand breaks. Presence of one type I and one type II topoisomerase is a minimal requirement for life but at least eight distinct families of topoisomerases exist in the biosphere. The emergence and the evolution of this plethora of topoisomerases is poorly understood. DNA gyrase is special type II topoisomerase with the unique capacity to introduce negative supercoiling in DNA. This enzyme is a major antibacterial target because of its essential role in homeostatic regulation of DNA supercoiling in nearly all bacteria. Intriguingly, despite its potential to interfere with essential DNA transactions, bacterial gyrase was successfully established in Archaeal domain by horizontal gene transfer but its emerging function and evolutionary history in this domain of life remain unclear. Here, we established archaea-specific and a global phylogeny of gyrases in prokaryotes. We find that the gyrase was introduced only once in the archaeal domain after the diversification of this group had already started. The archaea-centred phylogeny mostly recovers the major archaeal phyla suggesting mostly vertical evolutionary trajectory and only few early horizontal gene transfers within the archaeal domain. The global gyrase tree recapitulates the recently established deep divide between Terrabacteria and Gracilicutes suggesting that one of these superphyla was the gyrase donor. This topology further suggests that archaeal diversification started before the diversification of one of these two major bacterial groups.

INTRODUCTION

Topoisomerases are enzymes in charge of DNA topology regulation in every living cell. By their transient DNA cleaving activities, they maintain DNA supercoiling in a range compatible with DNA transactions such as transcription and DNA replication and ensure decatenation of chromosomes prior to cell division (Wang 2002; Schoeffler and Berger 2008; Pommier 2012; Bush, Evans-Roberts, and Maxwell 2015; Seol and Neuman 2016; McKie, Neuman, and Maxwell 2021). Topoisomerases are classified in type I and type II depending whether they cleave one or two DNA strands (Schoeffler and Berger 2008). In principle, only one type I and one type II topoisomerases would be necessary and sufficient for resolving all topological problems resulting from the DNA being very long molecule and those that build up during natural processed involving opening of the double helix. However, within these two classes, there are currently eight subfamilies of topoisomerases distributed unequally among the three domains of life indicating that the evolutionary history of these enzymes is complex. Known topoisomerases are all DNA-specific enzymes and their activity strongly affects DNA-templated processes in cells. Consequently, their emergence and evolution are intimately linked with the origin of DNA and its evolution as a carrier of genetic information.

As a rule, topoisomerases have the capacity to relax supercoiled DNA but only two of them can also convert relaxed DNA into supercoiled DNA. Reverse gyrase can supercoil DNA positively and this enzyme is, as yet, the only known hallmark of thermophily (Forterre 2002). Positively supercoiled DNA is more resistant to melting and therefore the particular activity of reverse gyrase is thought to help preserve the genome integrity at high temperatures (Kikuchi and Asai 1984). DNA gyrase, an antagonist of reverse gyrase, catalyses the ATP-dependent introduction of negative supercoils in constrained DNA molecules (Gellert et al. 1976). According to the twin supercoiled-domain model, DNA gyrase removes the positive supercoils which accumulate in front of transcribing RNA polymerase or during DNA replication thus allowing these essential processes to proceed (Liu and Wang 1987; Drlica 1992; Lal et al. 2016; Sutormin et al. 2018). In tandem with the topoisomerase I (Topo I), DNA gyrase regulates global supercoiling level in bacterial cells and even small deviations from optimal DNA topology are lethal for bacteria (Pruss, Manes, and Drlica 1982). Gyrase activity is, *via* the intracellular ATP/ADP ratio, directly linked with global state of cellular metabolism. In stressful conditions such as lack of nutrients, the ATP levels drop, DNA gyrase is less active and, consequently, the global supercoiling level of DNA is modified. In response, the expression of numerous genes (between 7 and up to 48 % of all genes) are simultaneously altered to allow quick adaptation to such unfavourable conditions (Dorman and Dorman 2016; Martis B et al. 2019; Bush, Evans-Roberts, and Maxwell 2015). The capacity of the DNA gyrase to quickly translate environmental stimuli into appropriate global transcriptional response

is believed to be a key feature providing the bacteria with capacity to adapt rapidly to changing environment (Rui and Tse-Dinh 2003; Dorman 2006).

Bacterial DNA gyrase A and B subunits assemble into A_2B_2 heterotetramer of approximately 370 kDa forming three major subunit interfaces, or gates, called N-gate, DNA-gate and C-gate. In order to perform its function, gyrase undergoes a sequence of conformational changes that consist of concerted gate openings, DNA cleavage, and DNA strand passage events (Schoeffler and Berger 2008; Soczek et al. 2018; Vanden Broeck et al. 2019). The DNA gate houses the catalytic tyrosine and cooperates with the TOPRIM fold to cleave DNA (McKay and Steitz 1981; Harrison and Aggarwal 1990; Berger et al. 1996; Aravind, Leipe, and Koonin 1998; Gajiwala and Burley 2000). The unique C-terminal DNA binding domain (CTD) carries the conserved GyrA – box motif, Q(R/K)RGG(R/K)G, which has been identified as hallmark feature of DNA gyrases (Reece and Maxwell 1991; Qi, Pei, and Grishin 2002; Corbett, Shultzaberger, and Berger 2004; Ward and Newton 1997). This motif is essential for chiral wrapping of DNA around gyrase enabling these enzymes to introduce negative supercoils in relaxed or positively supercoiled DNA (Lanz and Klostermeier 2012; Kramlinger and Hiasa 2006).

In contrast to bacterial orthologs which were extensively studied *in vitro*, only one archaeal gyrase, the one from *Thermoplasma acidophylum* has been biochemically characterised. This enzyme exhibited *in vitro* activities similar to that of bacterial homologs e.g. ATP-dependent supercoiling and ATP-independent decatenation and relaxation activities (Yamashiro and Yamagishi 2005). Early studies showed that gyrase had negative supercoiling activity *in vivo* and that this activity was essential in methanogens, halophiles and thermoacidophiles (Sioud, Possot, et al. 1988). Whether, as in bacteria, the gyrase activity, and through it, the supercoiling homeostasis, is linked to chromosomal gene expression regulation in Archaea is poorly understood. The gene encoding rhodopsin in extreme halophile *Haloferax* was strongly induced when gyrase activity was inhibited by novobiocin drug indicating that supercoiling-sensitive promoters may exist in this archaeon. In another halophile *Halobacterium sp*. the treatment with novobiocin resulted in deregulation of many genes including the stimulation of gyrase, topoisomerase VI (Topo VI) and Topo I expression indicating the involvement of these enzymes in regulating the DNA supercoiling levels in this organism (Tarasov et al. 2011).

Previous studies using the taxonomic sampling available at the time, identified the presence of gyrase encoding genes in all members of monophyletic Cluster II Euryarchaeota but not in other known archaeal phyla (Forterre et al. 2007; Raymann et al. 2014). Thus, at the time, it was proposed that gyrase was transferred from Bacteria to Archaea, presumably only once at the base of Cluster II Euryarchaeota (Raymann et al. 2014). This major group within Euryarchaeota superphylum contains 7 distinct classes of organisms with various lifestyles (methanogens, halophiles, XXX) and, with exception

of Archaeoglobi, these are all mesophilic archaea. The acquisition of the gyrase by the thermophilic ancestor of Cluster II Euryarchaeota was further proposed to have played a key role in adaptation of this lineage to mesophilic lifestyle (López-García et al. 2015).

With the availability of an expanded taxonomic sampling it is now clear that the gyrase is also found in the archaeal lineages that are distantly related to the Cluster II Euryarchaeota such as DPANN and Asgard archaea (Adam et al. 2017). This raises the question of the timing and the number of inter- or intra-domain horizontal gene transfer events as well as the nature of the associated gene-specific or genome-wide adaptations. Resolving these questions is important for understanding the evolution of the Archaea, and also for testing hypotheses about the DNA topology-driven regulation of gene expression in Archaea, which is still poorly understood. To address these questions, we sampled gyrases from all available archaeal and bacterial genomes and performed phylogenetic, comparative sequence and structural analysis.

We found gyrase in all lineages of Cluster II Euryarchaeota, and in several lineages of DPANN and Asgard archaea thus expanding the gyrase dataset within these recently discovered archaea. Gyrase is also present in one lineage within Cluster I Euryarchaeota. The archaeal tree topology suggests that the gyrase was transferred from Bacteria to the base of Cluster II Euryarchaeota and then spread within other lineages via secondary horizontal gene transfers. Interestingly, we could not detect gyrase in TACK clade suggesting an incompatibility between the physiology of TACK and the activity of gyrase. Importantly, we found that global gyrase tree exhibits a tripartite topology whereby the bacteria form two clades corresponding to Terrabacteria and Gracilicutes and Archaea are monophyletic. This suggests that gyrase was transferred before the diversification of Terrabacteria and Gracilicutes but after the diversification of Archaea. Such a scenario could imply that the diversification of Archaea predates the diversification of Terrabacteria and Gracilicutes. Remarkably, deep branches separating bacteria from Archaea are very short. This suggests that gyrase acquisition by archaea has not been accompanied by drastic changes in its sequence and, consequently, that archaeal gyrases share very similar functions with the bacterial ones.

MATERIALS AND METHODS

Generation of the DNA gyrase data set

GyrA and GyrB archaeal data sets were constructed by searching in the National Center for Biotechnology Information (NCBI) data bases, for *Escherichia coli* GyrA (EGT66353.1) and GyrB (AKE86808.1) homologs using the BlastP program (Altschul et al. 1997). Additional searches were performed in Helarchaoeta and Verstraetearchaeota genomes by tBlastN (Altschul et al. 1997). Only sequences containing between 450 and 1100 amino acids and sharing at least 25 % sequence identity and more than 30 % of query sequence coverage were selected. Similar to other proteins involved in DNA metabolism, gyrase is a well-known carrier of inteins (Novikova, Topilina, and Belfort 2014). To detect inteins, GyrA and GyrB datsets were aligned separately and visualised using Geneious 11.0.5 (Biomatters), allowing the identification of 23 GyrA and 61 GyrB intein containing sequences which were removed from the data set. Using this approach, as of 5th of March 2020, 801 archaeal GyrA and 724 archaeal GyrB were retrieved. To complete GyrA and GyrB data sets, Annotree database was searched using KEGG accession numbers for GyrA (K02469) and GyrB (K02470) (Mendler et al. 2019). The 25565 results for GyrA and 25597 results for GyrB were downloaded. To avoid contamination of the GyrA data set by topoisomerase IV (TopoIV) A subunits, sequences lacking the GyrA box motif Q(R/K)RGG(R/K)G were eliminated. Dataset complexity was reduced by removing sequences with more than 90% sequence similarity across 70% of the sequence length for archaeal sequences and 70% sequence similarity across 70% of the sequence length for bacterial sequences using MMseqs2 v11.e1a1c (Steinegger and Söding 2017). The final dataset contained 385 GyrA and 348 GyrB archaeal sequences and 571 GyrA and 633 GyrB bacterial sequences. The dataset is available upon request.

Analysis of sequence conservation

Bacterial and archaeal sequence datasets were compared using BLASTp all against all analysis (E-value <0.001). The median value for sequence identity and dispersion around this value was calculated using R 3.6.1 (https://www.r-project.org/) and plotted using ggplot2 package with geom_violin function (https://CRAN.R-project.org/package=ggplot2).

Phylogenetic tree construction

Sequence alignment and trimming were performed using Galaxy web platform (Afgan et al. 2018). GyrA and GyrB data sets were aligned with MAFFT v7.273.1 (Katoh and Standley 2013) using BLOSUM30 matrix. Selection of phylogenetically informative positions in alignments was done using BMGE v1.12 (Criscuolo and Gribaldo 2010) with the following parameters: BLOSUM30 matrix, sliding

windows size 3, minimum block size 1. In parallel, les strict trimming was performed with Noisy (Dress et al. 2008) using 0.6 cut-off value to test for artificial shortening of branch lengths (Tan et al. 2015). Trees were generated using IQ-TREE multicore v1.6.7 (Nguyen et al. 2015) on the Genotoul computing cluster (http://bioinfo.genotoul.fr/). Branch support was assessed using ultrafast bootstrap method (UFBoot, 1000 iterations) (Hoang et al. 2018) and Shimodaira-Hasegawa approximate likelihood ratio test (SH-aLRT, 1000 iterations) (Guindon et al. 2010). One would typically start to rely on the clade if its SH-aLRT >= 80% and UFboot >= 95%. Automatic sequence evolution model selection was performed using ModelFinder with –m MFP option (Kalyaanamoorthy et al. 2017). Tests of tree topology were performed in IQ-TREE with 10,000 resamplings using the RELL method (Hirohisa Kishino, Miyata, and Hasegawa 1990; Hasegawa and Kishino 1994).

Trees were visualized using FigTree v1.4.4 (http://tree.bio.ed.ac.uk/software/figtree/). We systematically investigated the genomic context of highly divergent sequences (appearing as long branches in the trees) or sequences branching outside their taxonomic group to detect contamination with misannotated or outlier fast evolving sequences. Sequences from small contigs (less than 10 kb) or from contigs carrying rRNA genes not congruent with their annotation were removed from the data set. The dataset used for building final phylogenetic trees are available upon request.

For building a concatenated GyrA_GyrB tree, we first confirmed the presence of both *gyrA* and *gyrB* gene within the same genomes using Taxonomy ID. The two proteins were aligned using MAFFT, concatenated, and we performed the phylogenetic analysis using IQ-tree with the MPF option for sequence evolution model selection.

Synteny analysis

The synteny conservation analysis around the *gyrAB* locus was performed using the Clinker tool (Gilchrist and Chooi 2021). We selected organisms for which we detected, based on our phylogenetic analysis, putative HGT of gyrase genes. We extracted the corresponding GenBank file containing the contig encoding the *gyrA* and or the *gyrB* genes or a 30kb window of the chromosome or contig (if > 50kb). Global comparative analysis of our selected set of genbank files was done using default parameters. The gene clusters were determined based on protein similarity and the color-coded graphics was plotted.

Analysis of gyrase distribution in Archaea

The archaeal phylogenetic tree generated from 122 core proteins was retrieved from the Genome Taxonomy Database (GTDB Release 03-RS86) (Parks et al. 2020). The presence of GyrA or GyrB proteins

was searched using AnnoTree (Mendler et al. 2019) with additional manual curation to eliminate false positives (TopoIV). The presence/absence profiles were visualized using iTOL (Letunic and Bork 2021).

***gyrBA* operon conservation analysis**

To assess how conserved is the *gyrBA* operonic structure in Bacteria and Archaea we performed systematic search using the synteny maps option grouped by taxonomy available on the Genome Browser of the MicroScope platform (Vallenet et al. 2020) (https://mage.genoscope.cns.fr/microscope.)

Each line refers to a taxon, followed by the number of species within the taxon and having the same gyrase gene organization. The taxonomic rank can be modified through the «Option» button. For each taxon, the *gyrA gyrB* gene organization is symbolized by a colored box. The color of the box corresponds to percentage of species which have the same organization comparted to the reference gene. This percentage is computed by dividing the number of species having the particular gene organization with the total number of species belonging to this taxon. For under-represented bacterial taxonomic ranks we used the synteny conservation tool SyntTax. The search was performed using GyrA and GyrB proteins of *E. coli* (Oberto 2013) .

RESULTS

Collection and analysis of gyrase sequences

The phenomenal amount of sequence data available in the public sequence databases represents a valuable asset for understanding the evolution and diversity of life using phylogenomics and comparative genomics approaches. We took advantage of this resource to establish a comprehensive comparative sequence analysis of the DNA gyrase in Archaea. As initial step in this analysis, we constituted a curated set of GyrA and GyrB sequences across entire phylogenetic diversity of Archaea and Bacteria (see materials and methods).

Briefly, GyrA and GyrB datasets were constituted by searching for *E. coli* GyrA or GyrB orthologs in NCBI nr database by protein BLAST and by searching Annotree (REF). Intein-containing sequences were identified using sequence alignment and removed from dataset. We also removed redundant sequences, fast evolving sequences and misassigned sequences from metagenome-assembled genomes. To avoid contamination with subunit A of TopoIV, the GyrA dataset was filtered for GyrA box motif. GyrA box motif (Q(R/K)RGG(R/K)G) is the hallmark of DNA gyrases and its mutation abolishes the negative supercoiling activity (but not relaxation and decatenation activity) (REF). The final dataset contained 385 GyrA and 348 GyrB archaeal sequences and 571 GyrA and 633 GyrB bacterial sequences. Archaeal GyrA sequences contained the classical GyrA Box motif suggesting that the negative supercoiling function was conserved in archaeal gyrases (**Figure 1, A**). Using pairwise BLAST searches, we found that the sequence identity between archaeal and bacterial orthologs is comparable to sequence identity within archaeal or bacterial sequences (**Figure 1, B**). Together, the data suggest low amount of sequence evolution since introduction of DNA gyrase into archaeal domain.

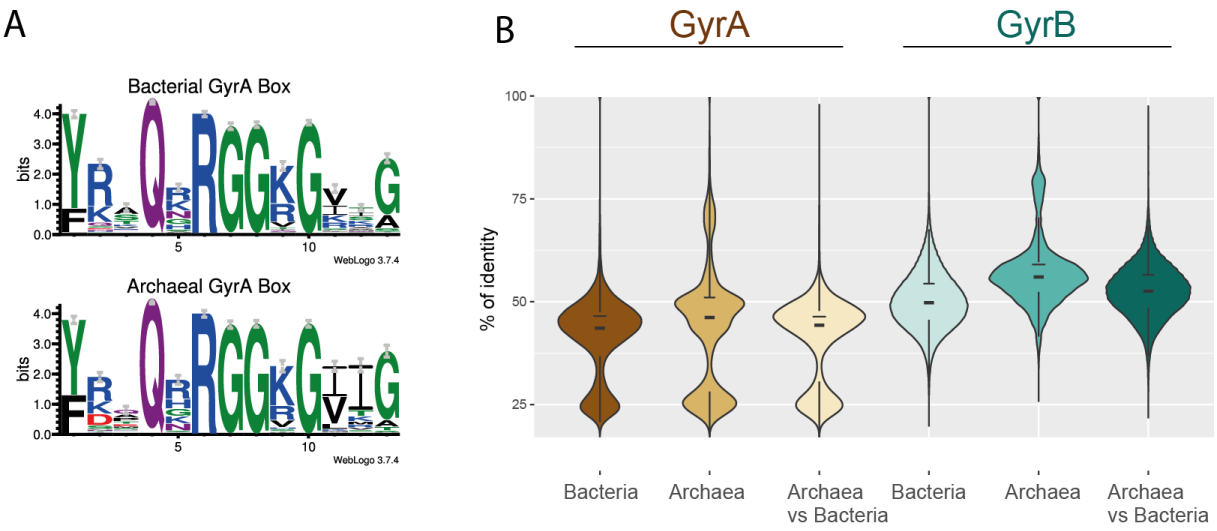


Figure 1. Sequence conservation of archaeal and bacterial gyrases

A) Comparison of the GyrA box motif. The motif was generated using an alignment of 477 bacterial and 299 archaeal representative sequences. The size of letters is proportional to the frequency of occurrence of each letter in the alignment. WebLogo v. 3.7.4. (REF) was used to generate the sequence logo.

B) Sequence conservation analysis using BLAST.

Sequence identity was determined using all against all BLASTp search. The statistical analysis and graphical representation were generated using R packages (see materials and methods). The analysis shows that sequence conservation within each domain or between the two domains is similar.

Distribution of gyrase in archaeal domain

Previous work showed that the gyrase is mainly present in one monophyletic group of Euryarchaea and, sporadically, in other archaeal lineages. We took advantage of the rapid increase of sequence data in public databases to update the gyrase distribution in Archaeal domain (**Figure2**).

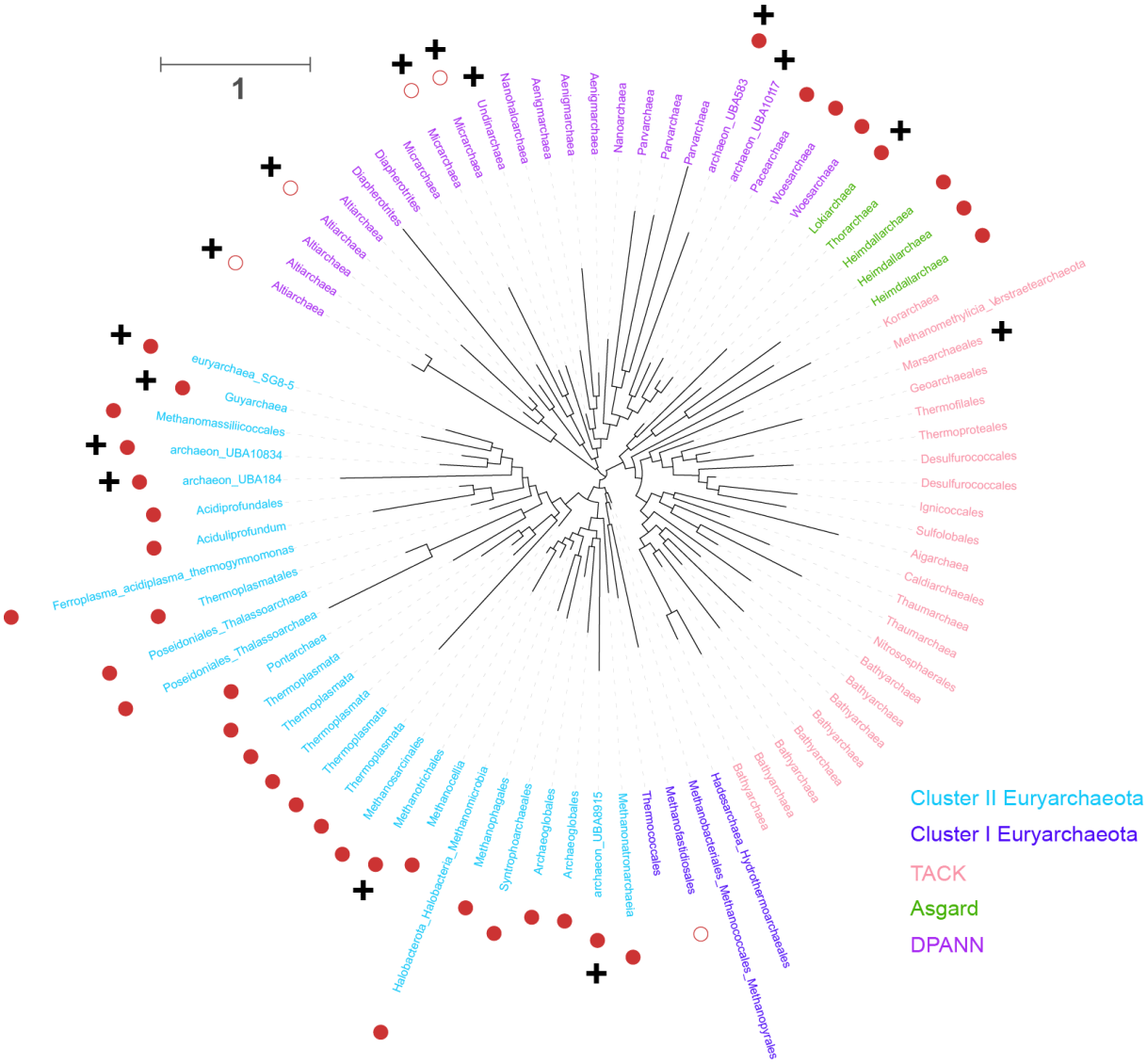


Figure 2. Distribution of gyrase across archaeal diversity.

The depicted phylogenetic tree shows the distribution of gyrase at the taxonomic rank of order. The presence of gyrase in the majority or all representatives is indicated with full red circle. The empty circle indicates presence in a few members only. It should be noted, however, that absence of genes in uncultured taxa may be due to genomes incompleteness. Plus sign indicates lineages that were discovered or significantly enriched in new genomes since the last survey of gyrase distribution by Adam and colleagues (Adam et al. 2017).

As one result of this survey, we confirmed the systematic presence of gyrase in Cluster II Euryarchaea and the systematic absence of gyrase in TACK superphylum including newly discovered lineages. The gyrase is also almost totally absent in Cluster I Euryarchaeota with exception of sporadic presence in Theionarchaea. We detected gyrase within DPANN superphylum in majority of Micrarchaeota, Woesearchaeota, Pacearchaeota, and in recently described (and still nameless) UBA583 order. Gyrase is only sporadically present in Altiarchaeota and is completely missing in the remaining seven DPANN

orders. In asgard archaea gyrase is robustly present in Loki- and Heimdall-archaeota but is missing in Thor- and Odin-archaeota. The diversity of Asgard archaea was recently substantially expanded by proposal of six additional phyla (Liu et al. 2021). Using BLAST search we detected gyrase genes in all phyla within a monophyletic clade composed of Heimdall-, Kari-, Gerdr-, Hod- and Wukong-archaeota and also in distantly related Loki- and Hela-archaeota. In contrast, gyrase genes were not detected in two new lineages, Baldr- and Hermo-archaeaota which form sister group with Odin- and Thor-archaeota, respectively.

The presence/absence pattern when superposed to the archaeal phylogenetic tree makes it difficult to deduce the evolutionary history leading to present-day distribution of the gyrase in archaea. Was gyrase present already in the last common ancestor of Archaea (LACA) and was subsequently lost in many lineages, of was it introduced later on as suggested by previous phylogenetic analyses? If the latter is true, how many intra or inter horizontal gene transfers occurred? Prompted by these questions we reconstructed the phylogeny of prokaryotic gyrases.

Archaeal and Bacterial gyrases segregate in three monophyletic clades

The above described distribution could be a result of vertical inheritance of gyrase genes from the last common ancestor of Archaea (LACA) or a consequence of a more recent inter- and/or intra- domain lateral gene transfers. The former scenario would imply that gyrase was already present in LUCA or that it was acquired by early lateral transfer from Bacteria before the diversification of Archaea. The latter scenario implies that the gyrase originated in bacterial lineage and was transferred to archaea after this domain diversified. These two scenarios lead to hypotheses testable by phylogenetic analyses: if the gyrase evolved in the lineage leading to the LACA, the global gyrase phylogeny should produce a bipartite tree where archaeal and bacterial sequences are monophyletic, whereas post-LACA inter-domain lateral spreading of gyrase would likely not produce a congruent bacterial and/or archaeal tree.

In order to investigate the evolutionary history of the gyrase we first built a phylogenetic tree using the alignment of concatenated GyrAB archaeal and bacterial sequences, respectively. The initial alignment was trimmed with BMGE to remove phylogenetically unreliable columns and increase the signal to noise ratio (Criscuolo and Gribaldo 2010). This filtering method removed 84 % of positions (1288 remaining positions out of 7978). In the phylogenetic tree inferred from this alignment, the bacterial sequences segregated into three separate clades while the archaeal sequences were split into two clades (**Supplementary Figure 1, A**). However, the ultrafast bootstrap and SH-aLRT values (36%/36%) did not support the existence of the central bipartition branch, suggesting that all archaeal

sequences may fall within a single monophyletic clade. This was further suggested by the tree containing only GyrA sequences (**Supplementary figure 6, A**) in which archaeal sequences formed a monophyletic clade with a fairly good branch support (97%/72%). To test this hypothesis, we modified the global gyrase phylogeny such that Archaea and Bacteria form monophyletic groups, and subjected this tree to various tests of phylogenetic tree selection (Kishino and Hasegawa 1989; Strimmer and Rambaut 2002; Shimodaira 2002; Shimodaira and Hasegawa 1999). These statistical analyses ask whether the tree consistent with the monophyly of archaea has a significantly worse-likelihood score than the maximum likelihood tree. These tests soundly confirmed the monophyletic tree topology (**Supplementary figure 3**).

We hypothesised that the conflicting, paraphyletic or monophyletic, topologies regarding the archaeal clade could be due to insufficient phylogenetic signal in our dataset. Indeed, we consistently observed that the SH-aLRT and UFBoot values have a tendency to disagree especially for deep nodes (UFBoot < 95 but SH-aLRT > 80). Such discrepancies can occur for datasets with week phylogenetic signal, that is, with insufficiently long and/or insufficiently diverged sequences (Guindon et al. 2010). To test if this is the case, we used Noisy (Dress et al. 2008) which has high tolerance for columns containing gaps thus allowing to potentially harvest substantial phylogenetic signal that may be contained in gap-rich regions (Dessimoz and Gil 2010; Tan et al. 2015). Trimming the alignment of the complete gyrase dataset with this method resulted in removal of 60 % of positions (3217 remaining positions out of 7978) thus more than doubling the number of alignment positions used for building the tree. The inferred maximum likelihood tree showed tripartite topology with archaeal sequences forming a monophyletic clade (**Figure 3, Supplementary figure 1, B**). Notably, the deep branches supporting the monophyly of the three clades were now robustly supported by both UFBoot and SH-aLRT suggesting that the relaxed trimming increased the signal to noise ratio in our dataset. However, despite the significant increase in the number of analysed positions the branches dividing archaeal and bacterial clades remained short indicating low level of sequence evolution since splitting of the archaeal lineage from bacteria.

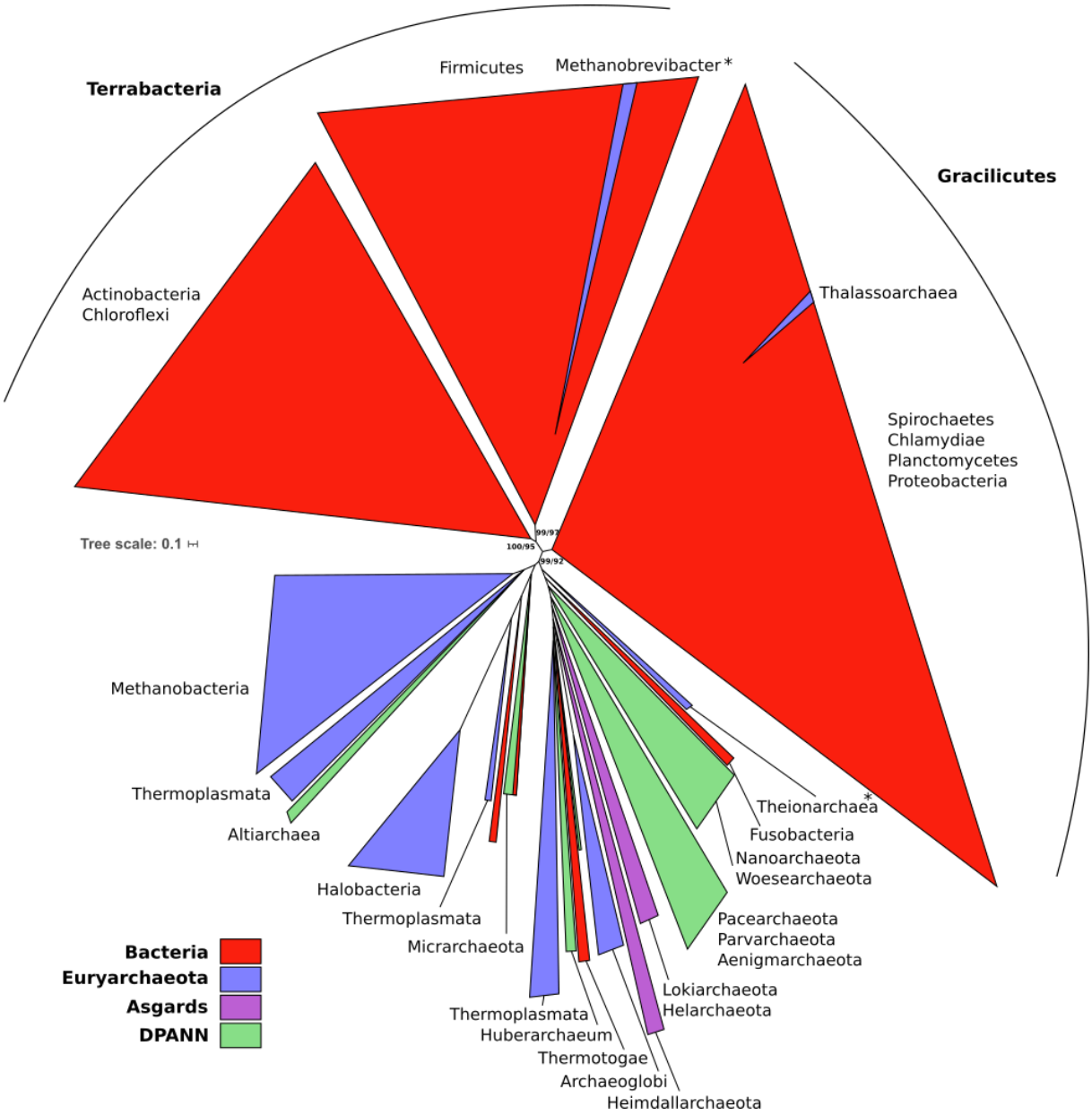


Figure 3: Global phylogeny of DNA gyrases

Maximum likelihood phylogenetic tree generated using the complete dataset of concatenated GyrA and GyrB sequences containing 502 bacterial and 297 archaeal sequences. The tree was inferred from an alignment trimmed with Noisy. The legend indicates the correspondence between the colours and taxonomic affiliations of the branches. SH-aLRT support (%)/ultrafast bootstrap support (%) for major bipartitions are indicated. Detailed tree is shown in **Supplementary figure 2, B**.

Remarkably, the bacterial part of the tree shows bimodal topology in agreement with the latest rooting of the bacterial tree (Coleman et al. 2021). The tree is divided into Terrabacteria (a clade encompassing Cyanobacteria, Deinococcus/Thermus, Firmicutes, etc.) and Gracilicutes (a clade encompassing Proteobacteria, Bacteroidetes, Spirochaetes, etc.) corresponding to the deepest phylogenetic divide between these two groups of bacteria suggesting mainly vertical evolution of gyrase in the bacterial

domain. Three alternative roots can be proposed based on this phylogenetic tree with different consequences for our understanding of the gyrase origin and evolution of bacteria and archaea (**Figure 4**). Rooting the tree in-between the archaeal clade and the two bacterial clades places the origin of gyrase before LUCA and subsequent partial loss of gyrase genes in archaeal domain (**Figure 4**, scenario I). Placing the root on the branch of Gracilicutes or of Terrabacteria leads to two different scenarii. Here, the gyrase originated in the branch leading to the last bacterial common ancestor (LBCA) and was then transferred to the lineage leading to Archaea (scenario II) or to the lineage leading to Cluster II Euryarchaeota with the donor bacterium belonging to the lineage giving rise to Gracilicutes or Terrabacteria. The scenario II implies that last archaeal common ancestor (LACA) is younger than LBCA while the third scenario infers that the transfer occurred between the donor bacterial lineage and the ancestor of Cluster II Euryarchaeota thus placing the origin of one of the major bacterial groups after LACA.

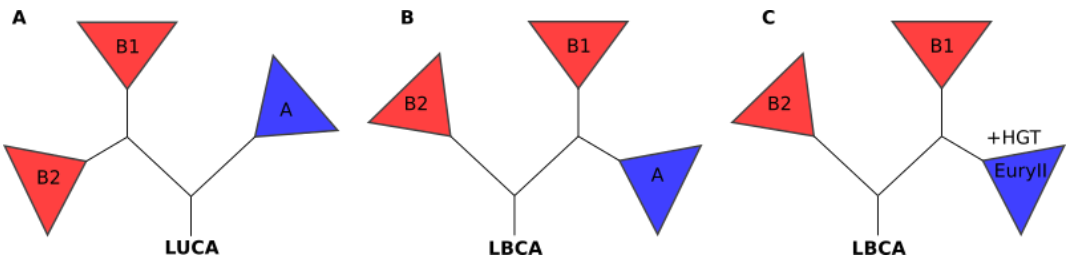


Figure 4. Evolutionary scenarii derived from the phylogeny of prokaryotic DNA gyrases

Assuming that the gyrase evolutionary history recapitulates faithfully, through mostly vertical inheritance of gyrase genes, the evolutionary history of organisms, three evolutionary scenarii can be deduced with consequences for the gyrase emergence and relative timing of emergence of major taxonomic groups.

A) The gyrase was already present in LUCA and LACA is older than LBCA. This scenario suggests that LUCA had DNA genome and that massive loss of gyrase genes occurred independently in the majority of archaeal lineages.

B) Gyrase emerged in bacterial lineage and LACA is younger than LBCA. This scenario suggests again occurrence of a massive and independent loss of gyrase genes in the majority of archaeal lineages.

C) Gyrase emerged in bacterial lineage and LACA is younger than the ancestor of Gracilicutes or Terrabacteria. In this scenario, the lineage giving rise to Gracilicutes (or Terrabacteria) is the donor of the gyrase and the recipient is the ancestor of Cluster II Euryarchaeota. The monophyly of Archaea can then be explained by spread of gyrase genes via secondary HGT within archaeal domain.

Collectively, the data show that the bacterial and archaeal clades are consistently separated by short branches indicating that low amount of evolution occurred in gyrase sequences since divergence from bacteria. This is inconsistent with the presence of the gyrase in the LUCA or its acquisition early in the

lineage leading to LACA because such history would most probably result in highly divergent proteins. Instead, our data agree with the idea that the introduction of the gyrase in the archaeal domain occurred once the tempo of evolution decreased i.e. after the diversification of this domain was initiated. Furthermore, the data suggest that the gyrase was introduced in archaeal lineage by a single horizontal gene transfer from bacteria and we propose that the ancestor of Cluster II Euryarchaeota was the recipient cell. The separation of bacterial sequences into two clades as distinct from each other as they are from archaeal clade suggests that the ancestor of Gracilicutes or Terrabacteria was the donor of the gyrase.

Phylogeny of archaeal gyrases

The global gyrase tree indicated that the archaea were monophyletic but internal topology of the archaeal clade was not robustly resolved. We reasoned that limited early HGT followed mostly by vertical inheritance would produce a phylogeny in which the majority of phylum-level taxonomic groups would be monophyletic, while frequent and recent HGT would likely not produce a typical archaeal tree. To test this hypothesis, we generated gyrase data sets containing 377 GyrA and 331 GyrB sequences, aligned them with MAFFT and performed trimming with BMGE or with Noisy. The unrooted archaeal phylogeny resulting from this analysis resolves into well supported clades corresponding to major taxonomic groups consistent with the consensus archaeal phylogeny but the three superphyla DPANN, Asgard and Euryarchaeota are not monophyletic in most of the trees (**Supplementary figure 4, 5 and 6**). More relaxed sequence alignment trimming resulted in GyrA phylogeny in which the Asgard monophyly was recovered and the DPANN were distributed between three closely branching clades (**Figure 5, A**). These two superphyla further branched within the Cluster II Euryarchaeota such that this group was split into two clades, one more closely related to Asgard archaea and the other to DPANN. In few cases, DPANN sequences branch within Asgard or Thermoplasmata clades suggesting recent HGT events or perhaps artefactual attraction of these fast-evolving DPANN sequences. Gyrases are almost completely absent in group I Euryarchaeota with exception of sporadic appearance in Theinoarchaea and Methanobrevibacter. The former organisms form a monophyletic clade branching within the group II Euryarchaeota and thus likely have acquired the gyrase by rather recent lateral transfer. The position of Methanobrevibacter within Asgard clade is most likely result of artefactual long branch attraction since Methanobrevibacter gyrases were recently acquired from bacteria (see below).

Although the use of relaxed trimming of sequence alignments recovers a tree topology more congruent with the expected archaeal tree topology, branch support at some deep nodes is systematically weak thus resulting at low resolution of the gyrase tree at inter-phylum level. These results are further

difficult to interpret since there is no consensus to the rooting of the archaeal tree (Petitjean et al. 2014; Raymann, Brochier-Armanet, and Gribaldo 2015; Da Cunha et al. 2017; Williams et al. 2017; Baker et al. 2020) and thus no true reference tree to compare with. Nevertheless, the repeated recovery of monophyletic clades corresponding to major taxonomic divisions within archaeal superphyla is not consistent with spread of the gyrase through multiple temporally separated HGT events. Rather, the data suggest predominantly vertical evolution with few early transfer events occurring before the diversification of archaeal superphyla. Such evolution fits with aforementioned evolutionary scenario III (**Figure 4**) suggesting that the gyrase was originally acquired by the lineage leading to group II Euryarchaea (which all encode gyrase) and later on introduced in the Asgard and DPANN superphyla.

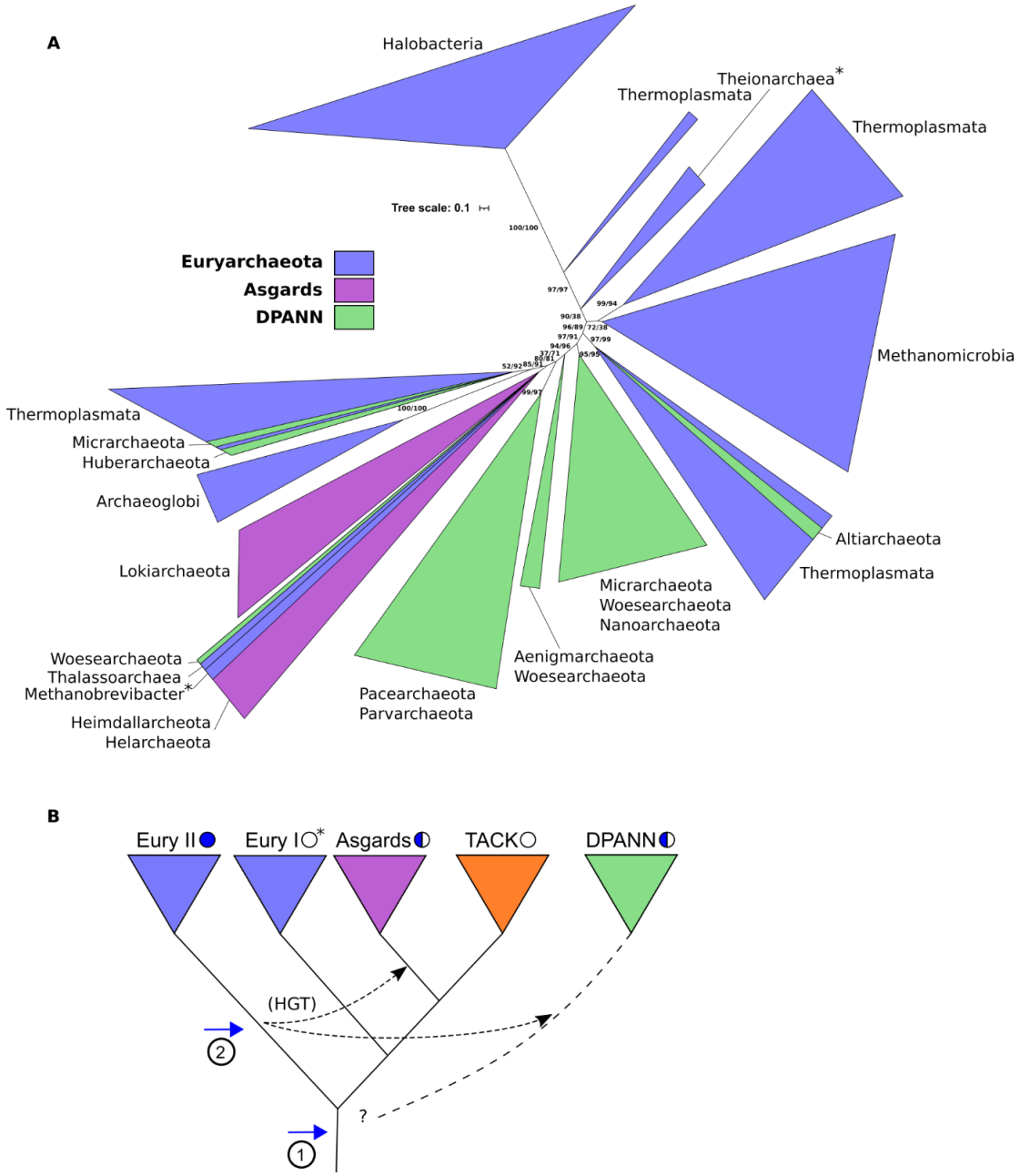


Figure 5. Vertical evolution is the predominant mechanism of gyrase spreading in archaeal domain.

A) Schematic representation of the maximum likelihood phylogeny of archaeal GyrA sequences. The detailed tree is shown in **supplementary figure 4**. The tree was inferred from the alignment of 376 GyrA sequences. The UFBoot support and SH-aLRT support are indicated for major bipartitions. Group I Euryarchaeota are highlighted with an asterisk.

B) Hypothetical scenario for gyrase evolution in Archaea. The acquisition of the gyrase by single HGT from bacteria before (1) or after (2) LACA is indicated with blue arrows. Putative HGT events toward ancestors of Asgard and DPANN superphyla are indicated with black arrows. The DPANN branch is represented by a dotted line to indicate the uncertain position of this group within the archaeal tree. The filled circles indicate the systematic presence of gyrase, semi-filled circles indicate partial presence and empty circles indicate absence of gyrase. Sporadic presence of gyrase in cluster I Euryarchaeota group is indicated by an asterisk. See main text for more detail.

How are the gyrase genes transferred horizontally between organisms?

Complete phylogeny of prokaryotic gyrases indicated that the inter-domain horizontal transfer of gyrase genes is uncommon thus suggesting that implantation of exogenous gyrase in a host cell from another domain is not trivial whether this cell already encodes an endogenous gyrase or not. Still, at least one successful implantation of bacterial gyrase in archaeal cellular context must have occurred early in evolution. Closer examination of cases where functional gyrase genes were recently transferred between archaea or from Archaea to Bacteria and *vice versa* could inform us about the necessary requirements for successful transfer.

When thinking about the possible mechanisms of gyrase spreading we first examined the possibility that gyrase could travel between organisms via mobile genetic elements. Indeed, for instance, the bacterial antibiotics resistance genes are notorious for being mobilised by plasmids, ICEs or viruses (MacLean and San Millan 2019). Recent survey of 38556 mobile genetic elements (ICE, plasmids and prophages) in Bacteria showed that *gyrA* genes are nearly absent in these elements indicating that their use as vehicles for horizontal spread of gyrase gene is, at best, very limited (Rodríguez-Beltrán et al. 2020). Using BLAST sequence similarity searches we performed similar survey across entire available plasmid and virus sequences of NCBI (44862 viruses and 31939 plasmids, Sept. 2021) and did not recover significant hits originating from plasmids whether they were from archaea or bacteria. This indicated that the gyrase genes do not figure among frequently transported genetic cargo carried by plasmids and suggest a different mechanism for horizontal transfer of gyrase genes in archaeal domain.

In case of viruses, we detected GyrA-like sequences in 13 isolated phages infecting *Bacillus* or *Lactococcus* bacteria but the closer inspection of these sequences revealed that they do not carry GyrA box motif suggesting that these are Topo IV subunits or degenerated DNA gyrases (**Supplementary figure 7, A**). We also detected hundreds of GyrA-like sequences in metagenome assembled phage genomes all originating from one recent meta-analysis of human-associated viruses (Tisza and Buck 2021). These metagenomes were classified as Microviridae, Myoviridae, Siphoviridae and Podoviridae all belonging to Caudovirales bacteriophages. The alignment of representative GyrA sequences for each of these taxa showed that Microviridae, Myoviridae and Siphoviridae contained the GyrA Box suggesting that these viruses may carry functional gyrases (**Supplementary figure 7, B**).

Collectively, the data show that mobile genetic elements rarely, if at all, encode gyrases except perhaps the bacteriophages from human microbiome. This suggest that the dispersion of the gyrase genes via these vehicles in not a common mechanism of transmission among organisms.

Cases of inter-domain horizontal gene transfer

To gain further insight in mode of transmission of gyrase genes we have examined the cases of interdomain HGT which were the only ones we could robustly detect. Indeed, in global gyrase phylogenies we consistently recovered the clustering of Methanobrevibacter archaea within Firmicutes bacteria while Thermotoga bacteria were positioned as a sister group to Huberarchaea from DPANN superphylum. Synteny analysis revealed that *gyrA* and *gyrB* genes are encoded on distant loci in Thermotoga bacteria (**Figure 6**). The synteny is supported by tree topology of Thermotogales thus suggesting vertical inheritance of gyrase genes from the common ancestor of this group. However, the genomic context is not well conserved across different genera suggesting a high degree of plasticity in this genomic region. Similar results were obtained for *gyrB* locus (**Supplementary figure 9**). We noted though that the gyrase containing 11-gene cluster assigned to *Kosmotoga olearia* was most similar to distantly related *Pseudothermotoga thermarum* indicating that this gene cluster may have been horizontally exchanged. In the most closely related archaeal donor species the gyrase genes are organised in a *gyrBA* operonic structure suggesting that this organisation was split into two different genomic locations in Thermotogales perhaps to facilitate the establishment of the archaeal gyrase in these bacteria.

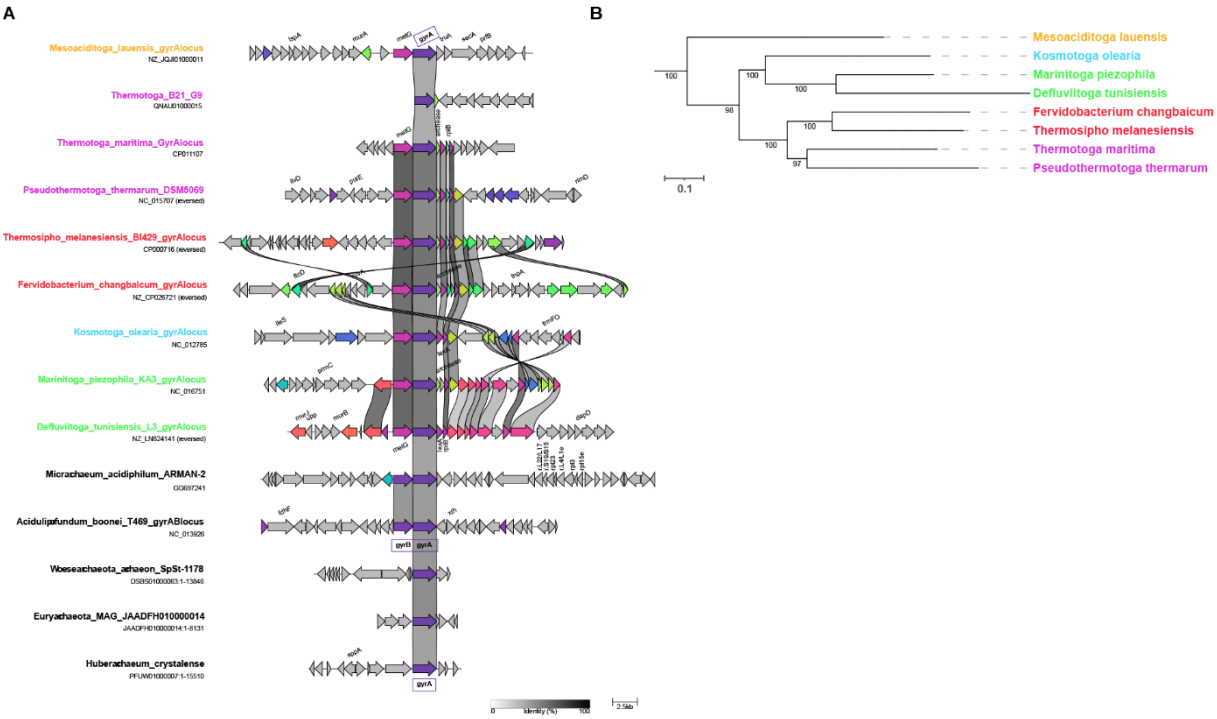


Figure 6. Synteny analysis of the *gyrA* locus in Thermotogales

A) The genomic context around *gyrA* gene is depicted with each arrow corresponding to a gene. Genes are automatically color coded based on functional annotation. The scale bar at the bottom corresponds to the percentage of identity between proteins encoded by the depicted genes. The drawing is on scale

with the scale bar representing 2.5 kbp. The bacterial species names are indicated in color and the same color code is used for the phylogenetic tree shown on the right.

B) Phylogeny of Thermotogales species used in synteny analysis. The phylogenetic tree was automatically generated using PhyloT and the Genome Taxonomy Database and visualized using iTOL (Letunic and Bork 2021). The numbers on the branches correspond to bootstrap node support values.

The transfer of gyrase genes from Firmicutes bacteria to Methanobrevibacter archaea must have been a relatively recent one since only few species of Methanobrevibacter carry gyrase. As such, this case can inform us about how gyrase can become established in a naïve organism. Synteny analysis showed that gyrase genes were clustered in *gyrBA* operon with no conservation of genomic context in vicinity not even between the two closely related species *M. curvatus* and *M. cuticularis* (**Figure 7**). This suggested that the gyrase operon may be mobilized as a stand-alone unit. Inactivation of gyrase genes by genetic drift would be expected if the gyrase offers no selective advantage to Methanobrevibacter species. Using sequence alignment, we detected all catalytically important residues in archaeal orthologs suggesting that these may indeed be functional (**Supplementary figure 10**). A single mutation with potential to modify the negative supercoiling activity was detected in GyrA box motif where one of the universally conserved glycines was mutated to alanine. The 170 amino acid insertion in the metal- and DNA-binding TOPRIM domain that occurs in the GyrB subunit of many gram negative bacteria (*including E. coli*) is missing in Methanobrevibacter species in line with these gyrases being of gram positive origin as suggested by the tree topology and synteny analysis.

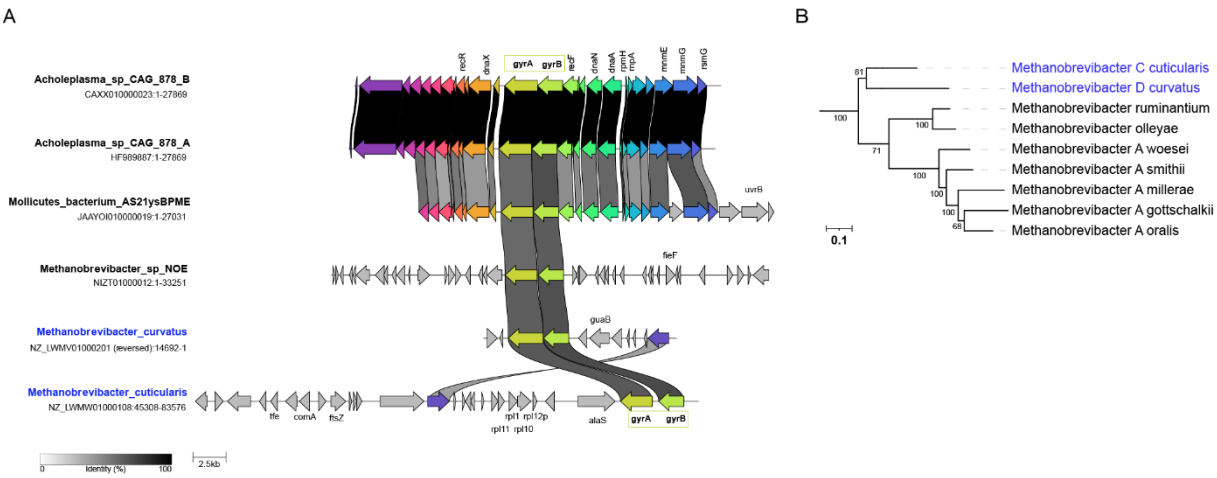


Figure 7. Synteny analysis of the *gyrBA* locus in Methanobrevibacter species

A) The genomic context around *gyrBA* operon is depicted with each arrow corresponding to a gene. Genes are automatically color coded based on functional annotation. The scale bar at the bottom corresponds to the percentage of identity between proteins encoded by the depicted genes. The drawing is on scale with the scale bar representing 2.5 kbp. The two closely related Methanobrevibacter species are highlighted in blue.

B) Phylogeny of Methanobrevibacter species. The phylogenetic tree was automatically generated using PhyloT and the Genome Taxonomy Database and visualized using iTOL (Letunic and Bork 2021). The numbers on the branches correspond to bootstrap node support values.

DISCUSSION

Topoisomerases are ancient enzymes with a complex evolutionary history. These ubiquitous enzymes are important for genome stability, replication and gene expression. As such, they must have played an important but still poorly understood role in the evolution of genomes. In this study, we focused on DNA gyrase which are ubiquitous in Bacteria but have a patchy distribution in Archaea. This topoisomerase maintains a tight control of chromosomal supercoiling in Bacteria and through it exercises a topology-driven control of gene expression. From evolutionary standpoint, the gyrase is of much interest because of its capacity to trigger massive gene deregulation in response to environmental stimuli. For this reason, it was suggested that gyrase is a key enzyme allowing the bacteria to adapt and colonise a variety of ecological niches. Much less is known about the *in vivo* function of Archaeal gyrases and about the role this enzyme played in the evolution of this domain of life. As initial approach to close this knowledge gap, we have analysed the phylogeny of prokaryotic gyrases with special focus on evolutionary history of archaeal orthologs. We found gyrase orthologs in Euryarchaeota, Asgard and DPANN superphyla but not in TACK superphylum. The constructed phylogenetic trees revealed that Archaea acquired the gyrase from bacteria probably by a single HGT and that the gyrase genes were further spread within archaea by secondary transfers. We further show that the most parsimonious evolutionary scenario implies that the gyrase emerged in the lineage giving rise to bacteria and that the archaea diversified before one of the major groups of Bacteria. Finally, the mode of transmission of gyrase genes seems not to include mobile genetic elements.

When thinking about the evolution of the gyrase we wondered why this enzyme is not more widely spread in Archaea? Plasmids and viruses are notorious as disseminators, via horizontal gene transfer, of a large diversity of genes and this mechanism is important for adaptation and genetic diversification of Bacteria and Archaea (Polz, Alm, and Hanage 2013; Medini et al. 2005). Recently, the members of the type IIB topoisomerase family (topoisomerase VIII and Mini-A) were detected encoded on free or integrated conjugative plasmids in Bacteria and Archaea and our own search shows presence of TopoIV in isolated bacteriophages indicating that these topoisomerases can be horizontally transferred by the archaeal and bacterial mobilome (Takahashi et al. 2020). On the other hand, this seems not to be the case for genes encoding gyrase orthologs thus limiting the dissemination potential for gyrase genes. In bacteria, the gyrase is important for maintaining the supercoiling levels within a range compatible with DNA transactions (Cheung et al. 2003; Dorman and Dorman 2016; Peter et al. 2004; Sutormin et al. 2018). As a part of this homeostatic mechanism, the promotors of *gyrA* and *gyrB* genes are stimulated by DNA relaxation (Menzel and Gellert 1983; Straney, Krah, and Menzel 1994; Unniraman and Nagaraja 1999). It seems therefore plausible to think that the introduction, via horizontal gene transfer, into this

tightly controlled system of an additional functional gyrase may cause imbalance leading to fitness cost or lethality.

Still, through our phylogenetic analysis, we detected two robust cases of relatively recent transfer of gyrase genes between archaea and bacteria showing that despite potential negative consequences, the establishment of exogenous gyrases is fundamentally possible. In both detected cases, the putative donor organism encoded gyrase as *gyrBA* operon. This makes sense since the colocation of genes within an operon facilitates their collective translocation via horizontal gene transfer and both genes are required to form functional tetrameric A_2B_2 holoenzyme. In Thermotogales however, the operonic structure was broken down to two distant genomic loci and similar situation is found in many bacterial and archaeal lineages (data not shown) indicating, at least in these lineages, the absence of selective pressure for the maintenance of operon structure. Since organism bearing clustered genes are more likely to act as successful donors (Lawrence and Roth 1996), the separation of gyrase genes into two distant genetic loci would probably limit horizontal spread among bacterial and archaeal genomes.

We also noticed little or no conservation of genetic context even between closely related organisms such as the two Methanobrevibacter species suggesting that the gyrase operon may be transferred as a stand alone unit rather than as a part of a mobile cluster of genes but the precise mechanism remains to be determined.

Interestingly, the gyrase genes found in Methanobrevibacter species probably encode functional gyrases suggesting that these Cluster I Euryarchaea archaea which naturally do not encode gyrase have successfully adapted to the presence of a pervasive negative supercoiling activity. How come such, *a priori*, toxic activity with a wide ranging consequences for essential DNA transactions can be so easily accommodated and adopted? We have recently demonstrated that, unexpectedly, introduction of bacterial gyrase into *Thermococcus kodakarensis* is well tolerated by this hyperthermophilic archaeon (Villain et al. 2021). Notably, the *T. kodakarensis* cells did not overexpress endogenous topoisomerases to counteract the excess of negative supercoiling, instead we proposed that the plectonemic negative supercoils can be efficiently absorbed by wrapping DNA positively around nucleosomes. This hypothesis is also supported by the distribution of histone genes in archaea which largely overlaps with gyrase distribution pattern (Adam et al. 2017) thus indicating that histones may be a key factor facilitating the establishment of gyrase in archaeal cells. Actually, as a result of our survey we noticed that all archaeal gyrase bearers also encode histones. Gyrase-encoding Thermoplasma archaea are, at first glance, an interesting exception to this rule because they lost the original archaeal histone genes but have acquired instead a histone analog (dubbed HTa) of bacterial origin (Hocher et al. 2019). In these cells, as additional peculiarity, the Topo VI, which is the main archaeal type II topoisomerase, has

equally been lost, leaving the gyrase as the sole type II topoisomerase probably in charge of chromosome decatenation and supercoiling homeostasis. As another result of our survey of gyrase gene distribution, we noticed that among ten thermophilic archaeal lineages (Catchpole and Forterre 2019), Archaeoglobi are the only ones carrying both gyrase and reverse gyrase, a topoisomerase with opposite, positive supercoiling activity and the sole known marker of hyperthermophily (Forterre 2002; Kikuchi and Asai 1984; Forterre et al. 1985). This observation indicates that the concomitant presence of these two enzymes with antagonist activities may be difficult to balance and is counterselected in Archaea. Interestingly, the only known example of a bacterium without functional gyrase is *Aquifex aeolicus*, one of the most extreme hyperthermophilic bacteria in which the endogenous gyrase was transformed into Topo IV enzyme (Tretter, Lerman, and Berger 2010).

What would be an evolutionary advantage for adopting a DNA gyrase? Through its negative supercoiling activity, the gyrase would be expected to change the delicate balance between the stability and melting potential of the DNA (Lopez-Garcia, 1999, Lopez-Garcia et al., 2015). In Bacteria, negative supercoiling drastically reduces the energy cost necessary for DNA melting and thus promotes transcription initiation by RNA polymerase (Martis et al 2019). In hyperthermophiles, it is presumed that the high temperature would be the energy source for melting the relaxed or positively supercoiled DNA (Meyer et al., 2013). Previous work exploiting the correlation between sequence composition and optimal growth temperatures suggested that early Archaea were (hyper)thermophiles, with mesophily arising more recently in archaeal evolution (Groussin et al., 2013, Groussin et al., 2013). Building upon this work, Lopez-Garcia and colleagues suggested that the acquisition of the bacterial gyrase by the ancestor of cluster II Euryarchaeota was a key evolutionary event allowing the progressive adaptation to mesophilic lifestyle (Lopez-Garcia et al., 2015). Indeed, through its negative supercoiling activity the gyrase could have facilitated DNA melting thus allowing the essential DNA transactions to proceed at suboptimal temperatures.

The evidence gathered so far indicates strongly that, as in Bacteria, gyrase has an active and essential role in regulating DNA supercoiling in Archaea (Sioud, Possot, et al. 1988; Sioud, Baldacci, et al. 1988). Together, these observations indicate that gyrase became essential in Archaea once mechanisms dealing with DNA became dependent on negative supercoiling. Indeed, in course of co-evolution with the gyrase, proteins involved in DNA transactions may have taken advantage of negative supercoiling of DNA for recognition and activity. Hence, a certain number of gene-specific and/or genome wide adaptations may have occurred since the acquisition of gyrase by Archaea. For example, DNA gyrase and histones/nucleosomes (which occur in many Archaea), both produce negative supercoils, but their mode of action is fundamentally different. Histones constrain supercoils in a solenoidal form (like a phone cord) while the gyrase produces plectonemic supercoil conformation (like a figure eight).

Moreover, the negative supercoils produced by gyrase are ready to assist in separating the double strand helix whereas the histones need to be *released* from nucleosomes to harness the energy contained in the negative supercoil. It's seems therefore plausible to presume that acquisition of gyrase may have impacted the evolution of genes encoding histones and, possibly, their function. Thus, the comparison of histone borne DNA condensation mechanisms between the gyrase encoding and gyrase lacking archaea provides the unique opportunity to explore evolutionary forces underlying the structuration of genomes.

A striking observation coming from phylogeny of gyrases is that the bacterial and archaeal clades are consistently separated by short branches. Similarly, the phylogenetic reconstructions of another topoisomerase, reverse gyrase, from bacteria and archaea consistently produced trees with short branches (Catchpole and Forterre 2019). This is clearly different from what is observed in phylogenies of universal marker proteins which typically exhibit very long branch separating bacterial and archaeal clades (Coleman et al., 2021). Such long branch is compatible with a fast tempo of sequence evolution which presumably occurred between LUCA and the ancestors of bacteria and archaea resulting in very divergent proteins (Woese, 1998, Forterre 2006). Consequently, the gyrase phylogeny as constructed from our sequence dataset, is inconsistent with the presence of the gyrase in the LUCA or its acquisition early in the lineage leading to LACA. Instead, our data agree with the idea that the introduction of the gyrase in the archaeal domain occurred once the tempo of evolution decreased i.e. during the diversification of this domain. According to the most parsimonious scenario, the gyrase was transferred only once from Bacteria probably into the lineage giving rise to the cluster II Euryarchaea. Given that the archaeal gyrase tree roughly recapitulates the established phylogeny of archaea, we suggest that the gyrase mainly evolved vertically with only few lateral gene transfers that must have occurred before the diversification of DPANN and Asgard superphyla. Recent phylogenetic analyses using outgroup-free approach allowed placing the bacterial root between Terrabacteria and Gracilicutes, two major bacterial clades (Coleman et al. 2021). In line with this proposal we recovered the deep split between these two clades thus indicating that, similar to archaeal counterparts, bacterial gyrase evolution has a major vertical component.

The analysis of ancestral horizontal gene transfers can also provide insight into the temporal sequence of events during gyrase diversification, because donor lineages must be at least as old as recipients. The topology of the gyrase tree suggests that Cluster II euryarchaeota may have been one of the earliest, if not the first, emerging archaeal lineages. This has implications for the ongoing debate concerning the position of the archaeal root for which there is currently no consensus (Petitjean et al.

2014; Raymann et al. 2015; Da Cunha et al. 2017; Williams et al. 2017). Moreover, according to our phylogenetic reconstruction, the archaea started their diversification before one of the two major bacterial clades emerged. Importantly, the gyrase evolution is incompatible with the presence of this enzyme in LUCA thus suggesting an RNA-based LUCA cell.

ACKNOWLEDGMENTS

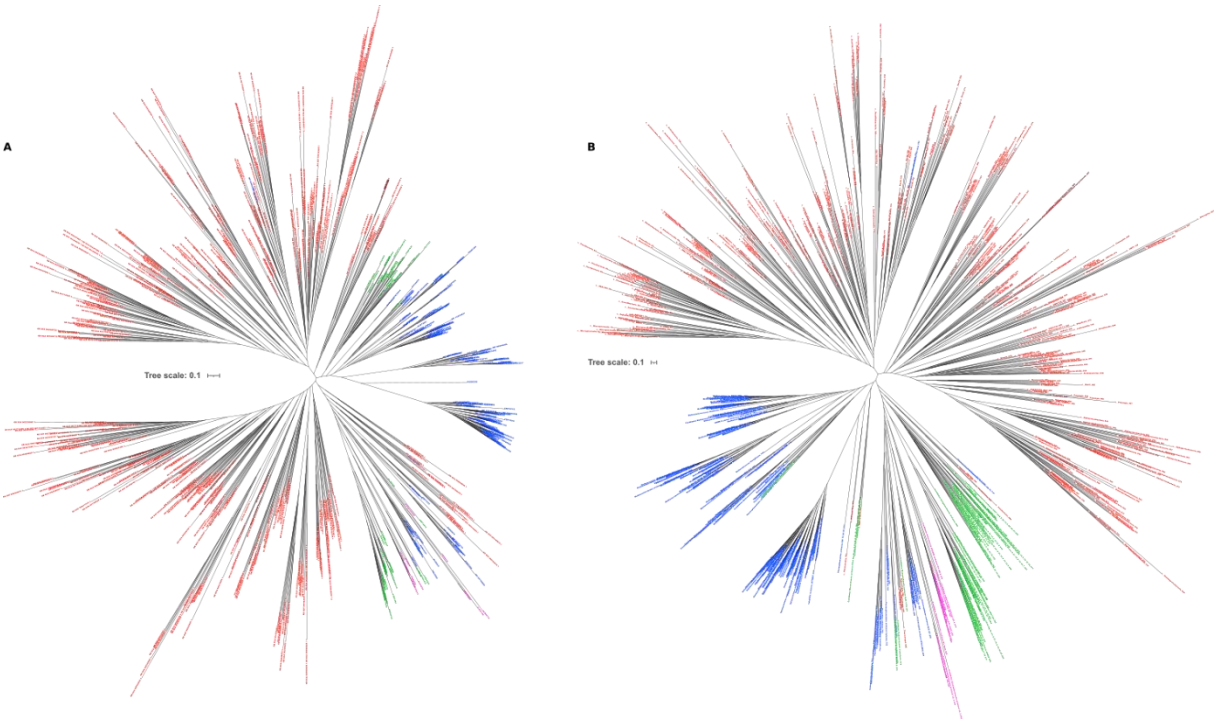
We are grateful to the Genotoul bioinformatics platform Toulouse Midi–Pyrenees for providing computing and storage resources.

SUPPLEMENTARY FIGURES AND TABLES

Supplementary table 1: Dispersion of sequence identity values for prokaryotic GyrA and GyrB sequences

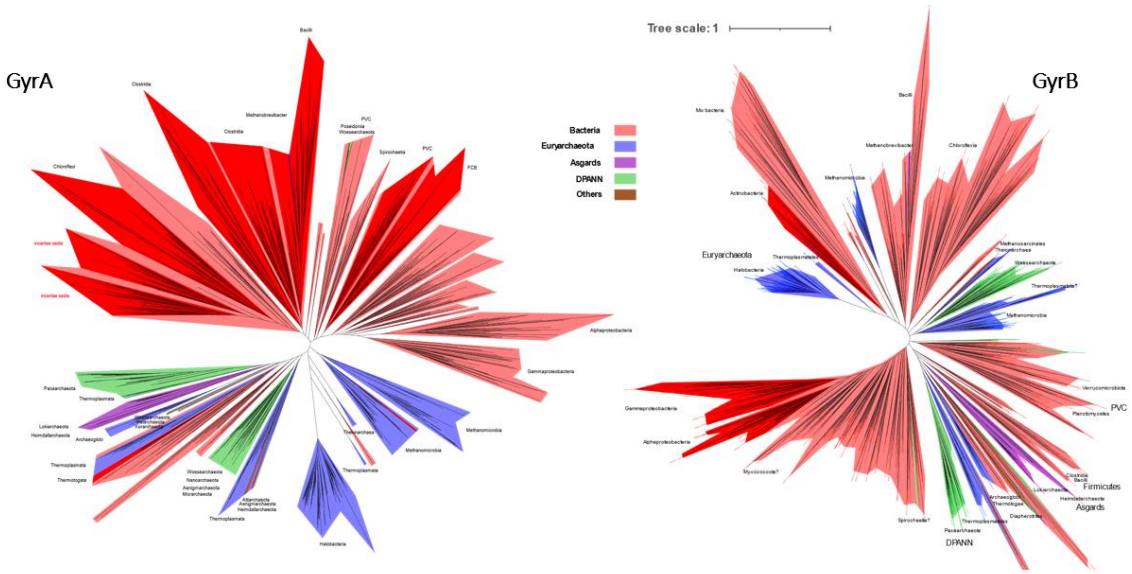
		Min.*	1st Qu.	Median	Mean	3rd Qu.	Max.
GyrA	Bacteria	16.67	36.75	43.55	41.11	47.40	100
	Archaea	16.79	28.14	46.17	43.67	51.02	100
	Archaea vs Bacteria	15.96	30.61	44.30	40.72	47.78	95.68
GyrB	Bacteria	21.12	45.57	49.75	50.21	54.36	100
	Archaea	27.18	52.38	55.99	57.25	59.62	100
	Archaea vs Bacteria	23.12	48.62	52.60	52.57	56.52	96.26

*All values are given as percentage



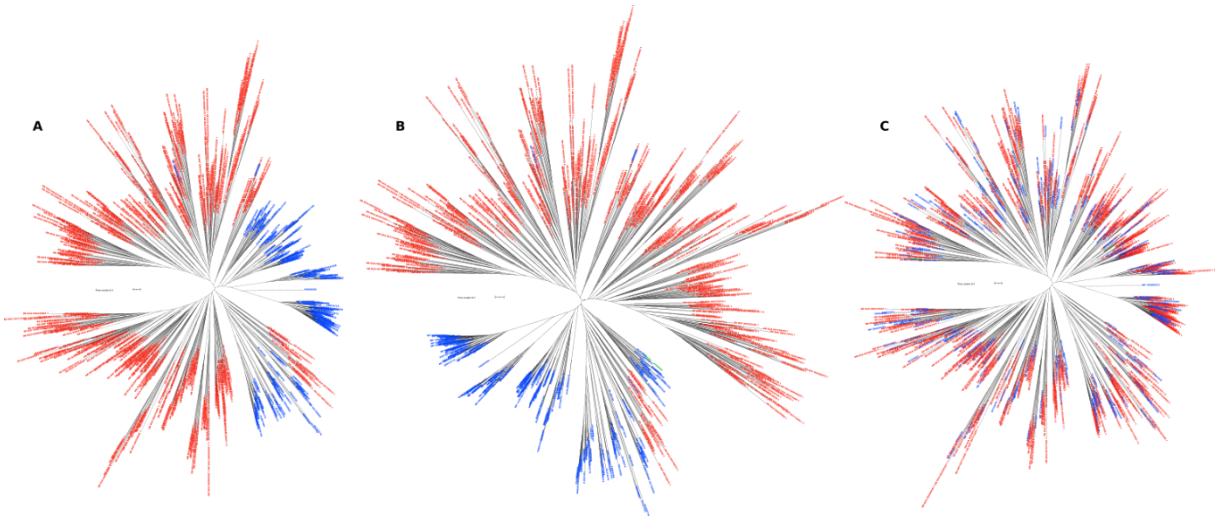
Supplementary figure 1. Global phylogeny of DNA gyrases

The GyrA and GyrB (502 bacterial and 297 archaeal) sequences were concatenated, aligned using MAFFT and trimmed using BMGE (A) or Noisy (B). This operation removed 6690 out of 7978 positions from the alignment when using BMGE and 4761 out of 7978 positions when using Noisy. Maximum-likelihood analysis was used for tree construction with IQ-TREE and branch support was assessed using UFboot and SH-aLRT (1000 iterations). The consensus trees are shown where sequences belonging to Euryarchaea, DPANN and Asgard superphyla are coloured in blue, green and magenta, respectively. The tree scale corresponds to number of substitutions per site.



Supplementary figure 2. Global phylogeny of GyrA or GyrB sequences inferred with stringent sequence alignment trimming

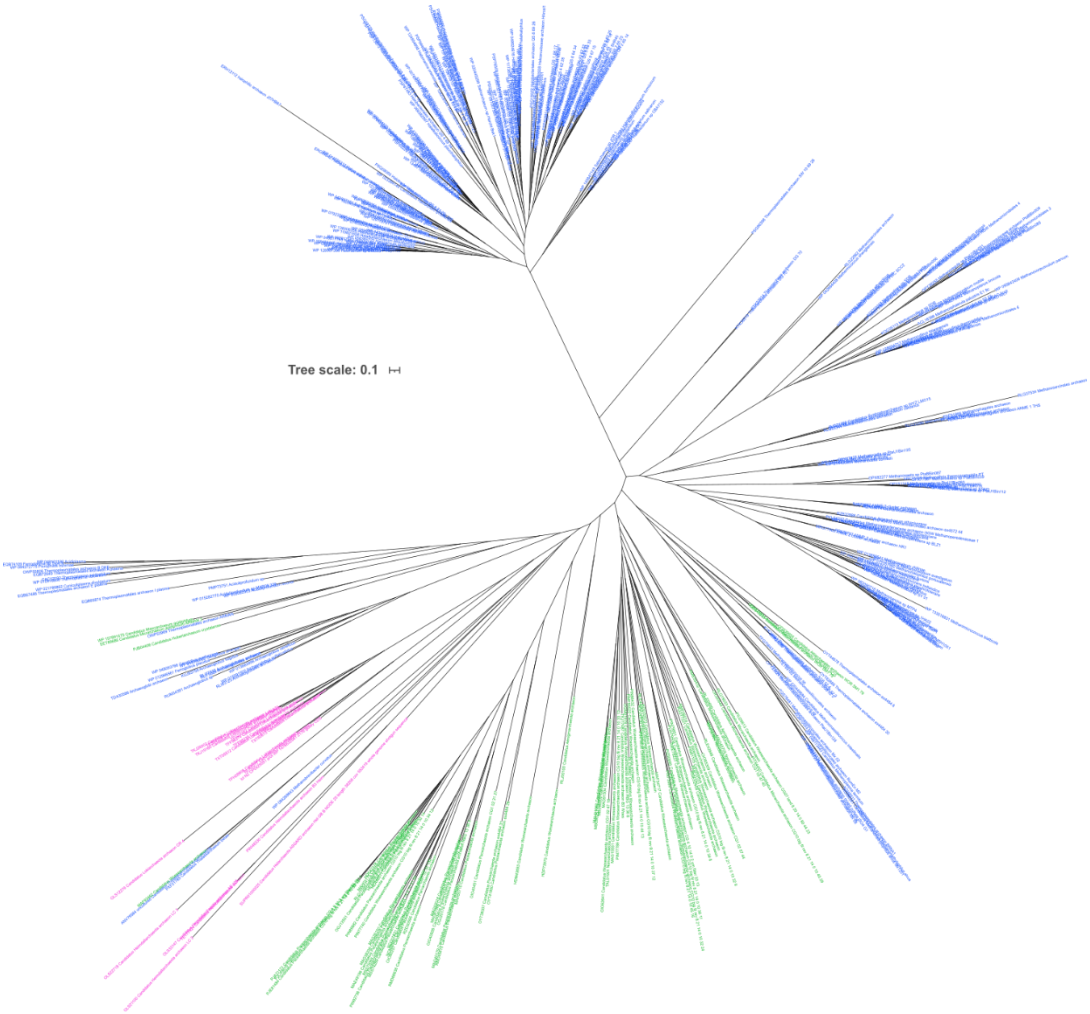
The GyrA (571 bacterial and 385 archaeal) or GyrB (633 bacterial and 348 archaeal) sequences were aligned using MAFFT and trimmed using BMGE. This operation removed 359 out of 1085 and 528 out of 1099 positions for GyrA and GyrB alignments, respectively. Maximum-likelihood analysis was used for tree construction with IQ-TREE and branch support was assessed using UFboot and SH-aLRT (1000 iterations). The consensus trees are shown where bacterial sequences are underlined in red. The legend indicates color-codes for the archaeal superphyla. The tree scale corresponding to number of substitutions per site is indicated.



	bp-RELL	p-KH	p-SH	p-WKH	p-WSH	c-ELW	p-AU
Tree A	0.551 +	0.554 +	1	0.554 +	0.802 +	0.549 +	0.561 +
Tree B	0.449 +	0.446 +	0.72 +	0.446 +	0.692 +	0.451 +	0.439 +
Tree C	0 -	0 -	0 -	0 -	0 -	0 -	2.61e-78 -

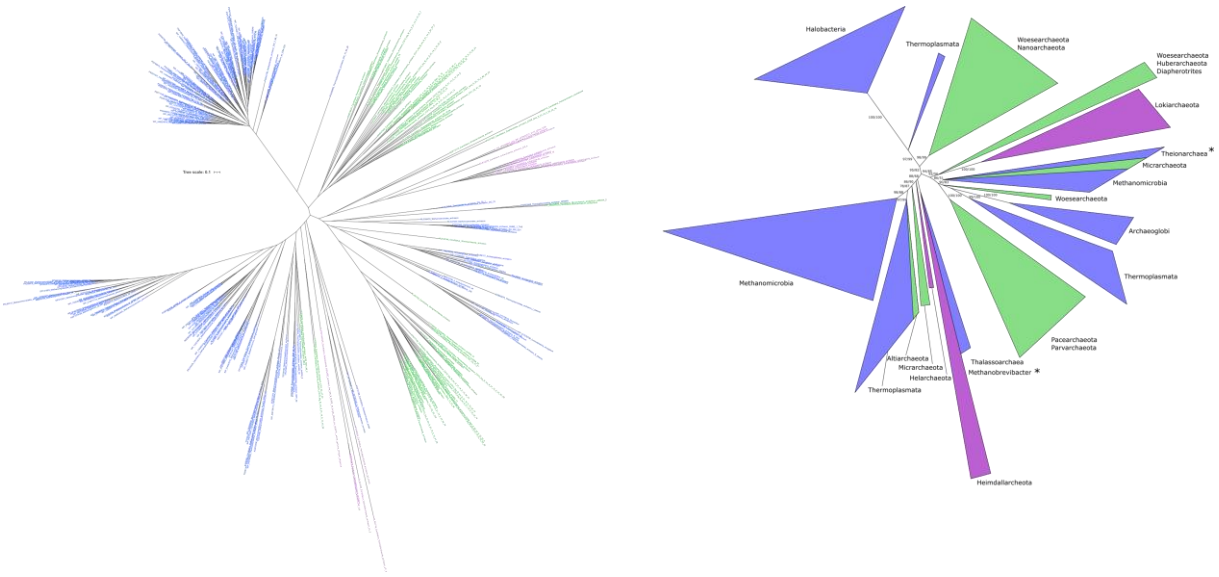
Supplementary figure 3. Global gyrase tree topologies used for tests of tree selection.

A) Complete gyrase phylogeny inferred from BMGE-trimmed alignment. B) The complete gyrase phylogeny where one deep branch was moved to produce monophyletic archaeal clade. C) Complete gyrase phylogeny where the branches were shuffled but the global tree topology was preserved. Results of tree selection tests are shown on the bottom part of the figure. Values indicate p-value for exclusion and are colored in green for tree topologies not excluded by the tests, red for tree topologies significantly excluded. KH is one-sided Kishino-Hasegawa test; SH is Shimodaira-Hasegawa test; WKH is weighted KH test; WSH is weighted SH test; ELW is Expected Likelihood Weight; AU is approximately unbiased test. Plus signs denote the 95% confidence sets. Minus signs denote significant exclusion. All tests were performed with 10000 resamplings using the RELL method.



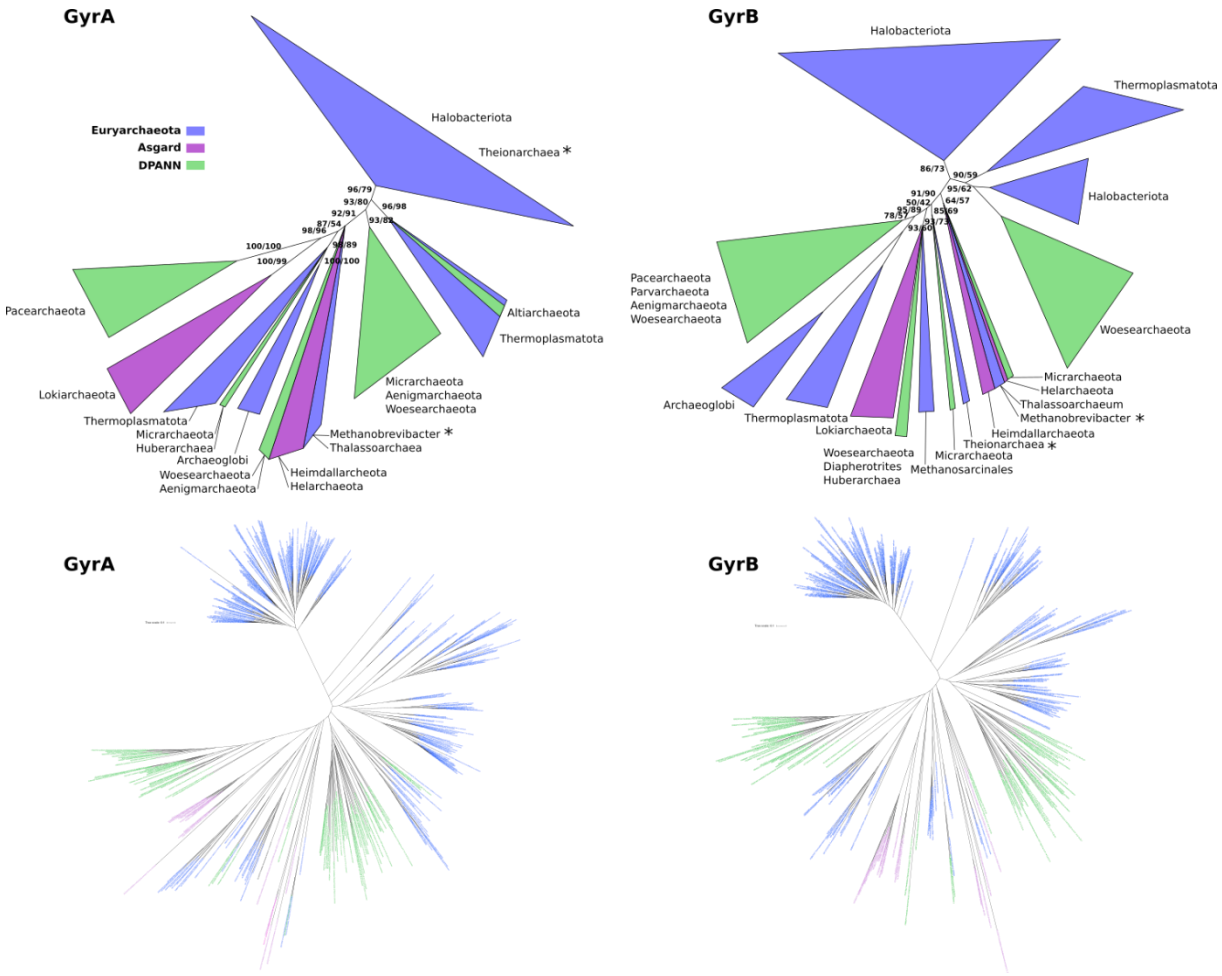
Supplementary figure 4. Maximum likelihood phylogeny of archaeal GyrA sequences.

Dataset containing 376 GyrA sequences was aligned using MAFFT and the phylogenetically uninformative regions were removed using Noisy. 1218 positions out of 1504 were used to produce the phylogenetic tree. The consensus trees is shown where sequences belonging to Euryarchaea, DPANN and Asgard superphyla are coloured in blue, green and magenta, respectively. The tree scale corresponding to number of substitutions per site is indicated.



Supplementary figure 5. Maximum likelihood phylogeny of archaeal GyrB sequences.

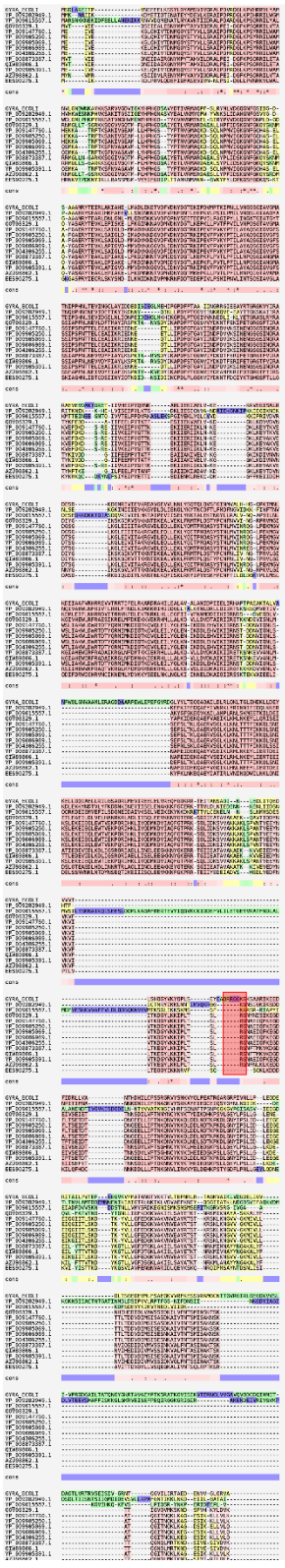
Dataset containing 331 GyrB sequences was aligned using MAFFT and the phylogenetically uninformative regions were removed using Noisy. 826 positions out of 1551 were used to infer the phylogenetic tree. Detailed consensus tree is shown on the left, sequences belonging to Euryarchaea, DPANN and Asgard superphyla are coloured in blue, green and magenta, respectively. The tree scale corresponding to number of substitutions per site is indicated. Schematic representation of GyrB tree is shown on the right. The members of group I Euryarchaeota are indicated with an asterisk.The numbers correspond to SH-aLRT and UFboot support for the deep nodes.



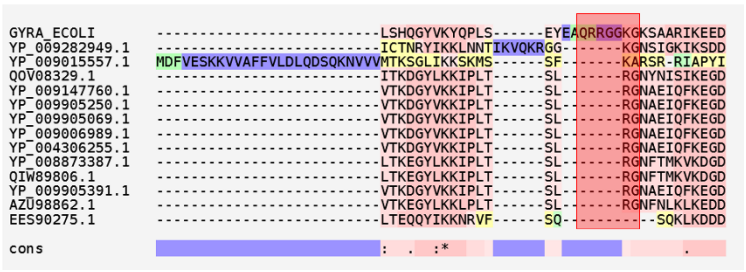
Supplementary Figure 6. Maximum likelihood phylogeny of archaeal GyrA or GyrB sequences inferred from conservatively trimmed sequence alignment.

377 GyrA and 331 GyrB sequences were aligned using MAFFT and trimmed using BMGE (see material and methods for details). The filtering removed 48 % of positions (779 from 1493 positions left) and 60% of positions (614 from 1554 left) for GyrA and GyrB alignment, respectively. Maximum-likelihood analysis was used for tree construction with IQ-TREE and branch support was assessed using UFboot and SH-aLRT (1000 iterations). The consensus trees are shown where sequences belonging to Euryarchaea, DPANN and Asgard superphyla are coloured in blue, green and magenta, respectively. The detailed trees are shown in the lower panel and their schematic representation is shown in the upper panel. The numbers on the branches indicate their UFboot and SH-aLRT support. Group I Euryarchaeota are indicated with an asterisk and the number of sequences present is indicated in brackets. The tree scale corresponding to number of substitutions per site is indicated.

A



B



Supplementary figure 7. Alignment of GyrA-like sequences from isolated bacteriophages.

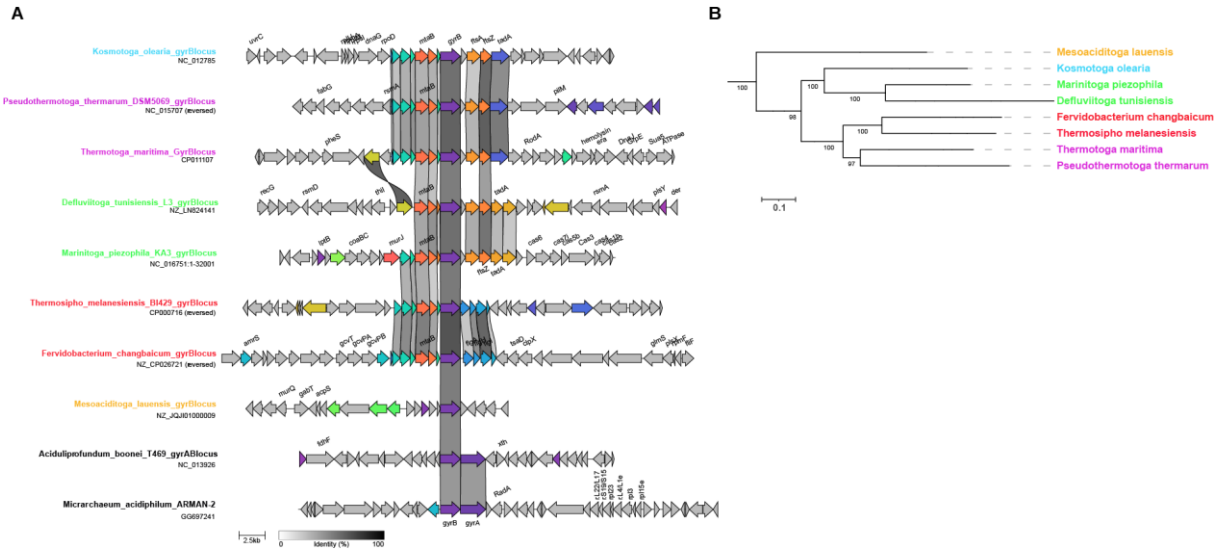
BLASTp search was used to identify viral homologs of *E. coli* GyrA (Uniprot ID: P0AES4). Viral sequences showing more than 25 % sequence identity over at least 70 % of sequence length were selected. Those sequences corresponded to bacteriohages infecting *Bacillus* or *Lactococcus* bacteria. Alignment was performed using T-coffee web server (Di Tommaso et al. 2011) with default settings. Part A shows the global alignment with GyrA box (QRRGGKG) highlighted with a red rectangle and part B is enlarged

part of the alignment containing the GyrA box. The first sequence is GyrA of *E. coli* which serves as a reference for canonical GyrA protein. The alignment shows that viral sequences do not contain GyrA box motif.

Supplementary figure 8. Alignment of GyrA-like sequences from metagenome assembled viral genomes

BLASTp search was used to identify viral homologs of *E. coli* GyrA (Uniprot ID: P0AES4). Viral sequences showing more than 25 % sequence identity over at least 70 % of sequence length were taken into consideration. This yielded over 100 sequences corresponding to Microviridae, Myoviridae, Siphoviridae and Podoviridae all belonging to Caudales bacteriophages. One or two sequences were selected for each taxonomic group for alignment which was performed using T-coffee web server (Di Tommaso et al. 2011) with default settings.

Part A shows the global alignment with GyrA box (QRRGGKG) highlighted with a red rectangle and part B is enlarged part of the alignment containing the GyrA box. The first sequence is GyrA of *E. coli* which serves as a reference for canonical GyrA protein. Four viral sequences possess the canonical GyrA box motif while in the remaining four sequences the motif is degenerated or missing.

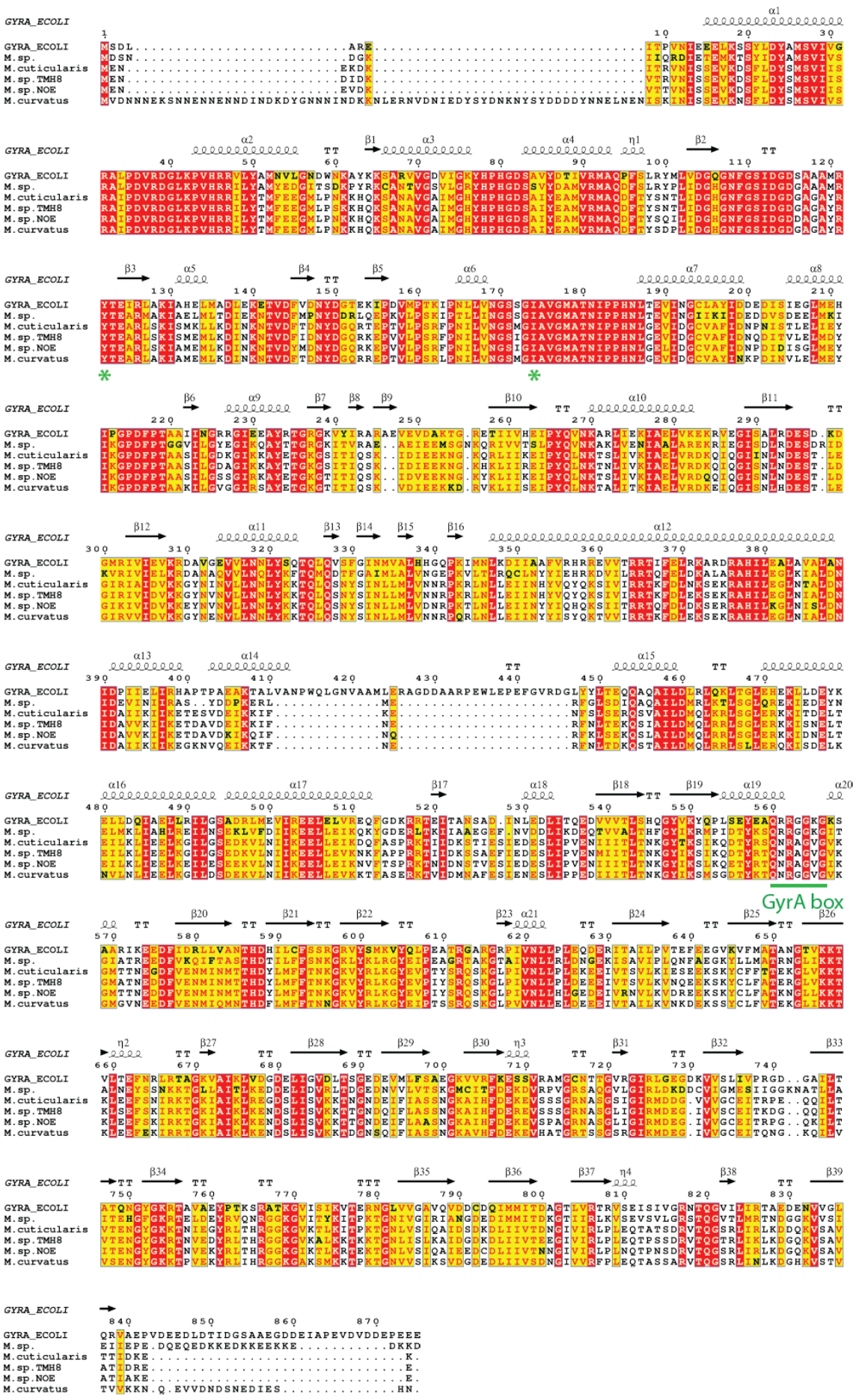


Supplementary figure 9. Synteny analysis of the *gyrB* locus in Thermotogales

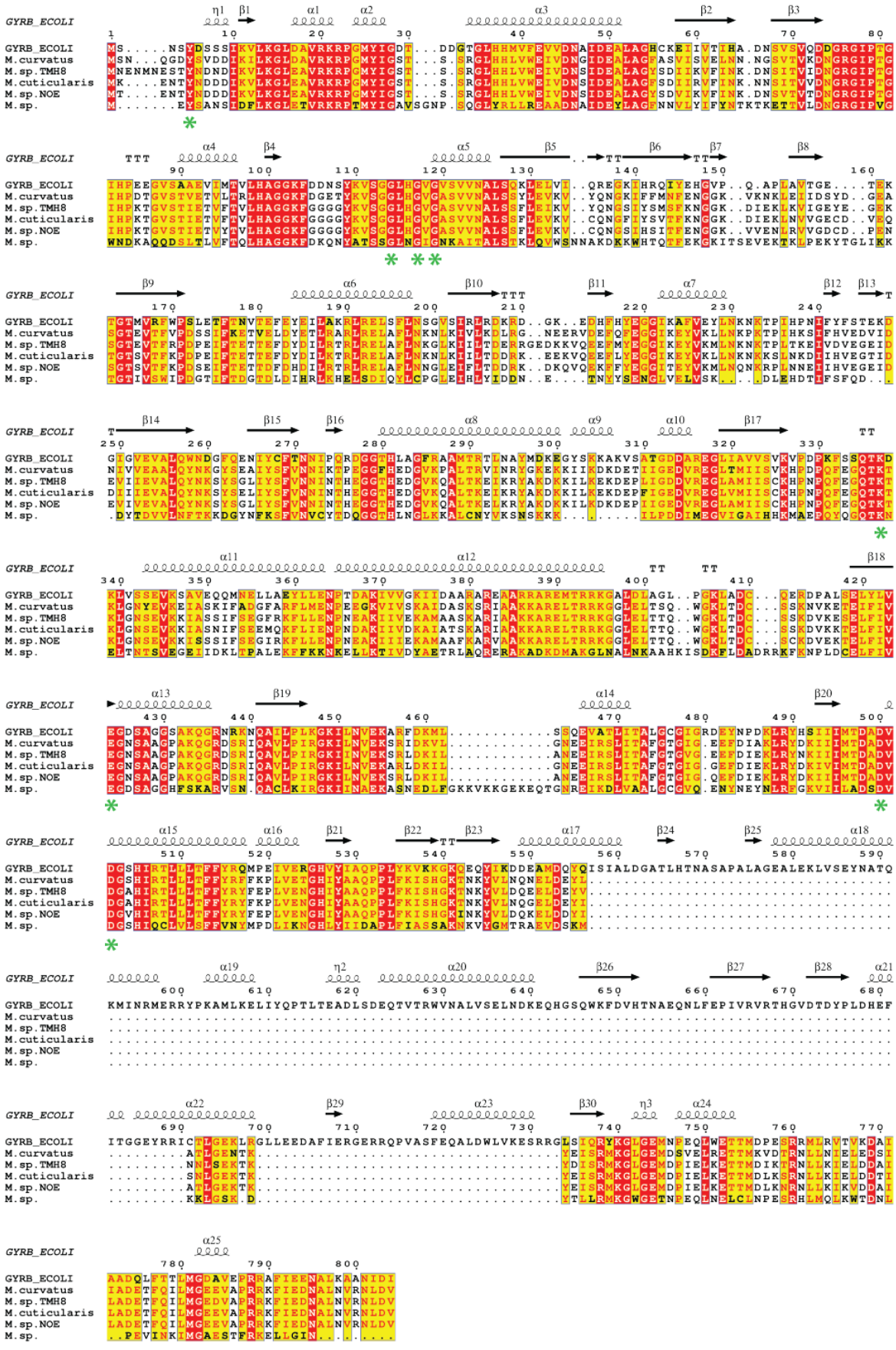
A) The genomic context around *gyrB* gene is depicted with each arrow corresponding to a gene. Genes are automatically color coded based on functional annotation. The scale bar at the bottom corresponds to the percentage of identity between proteins encoded by the depicted genes. The drawing is on scale with the scale bar representing 2.5 kbp. The bacterial species names are indicated in color and the same color code is used for the phylogenetic tree shown on the right.

B) Phylogeny of Thermotogales species used in synteny analysis. The phylogenetic tree was automatically generated using PhyloT and the Genome Taxonomy Database and visualized using iTOL (Letunic and Bork 2021). The numbers on the branches correspond to bootstrap values for node support.

A



B



Supplementary figure 10. Sequence alignment of GyrA and GyrB orthologs from Methanobrevibacter species.

Sequences were aligned using T-coffee web server (Di Tommaso et al. 2011) and secondary structure information was rendered using ESPript 3.0 (Robert and Gouet 2014). The catalytically important residues and the GyrA box are indicated with green star and green line, respectively. For full description of all catalytic residues, see Schoeffler and Berger, 2008.

REFERENCES

Adam, Panagiotis S., Guillaume Borrel, Céline Brochier-Armanet, and Simonetta Gribaldo. 2017. 'The Growing Tree of Archaea: New Perspectives on Their Diversity, Evolution and Ecology'. *The ISME Journal* 11 (11): 2407–25. https://doi.org/10.1038/ismej.2017.122.

Afgan, Enis, Dannon Baker, Bérénice Batut, Marius van den Beek, Dave Bouvier, Martin Cech, John Chilton, et al. 2018. 'The Galaxy Platform for Accessible, Reproducible and Collaborative Biomedical Analyses: 2018 Update'. *Nucleic Acids Research* 46 (W1): W537–44. https://doi.org/10.1093/nar/gky379.

Altschul, S. F., T. L. Madden, A. A. Schäffer, J. Zhang, Z. Zhang, W. Miller, and D. J. Lipman. 1997. 'Gapped BLAST and PSI-BLAST: A New Generation of Protein Database Search Programs'. *Nucleic Acids Research* 25 (17): 3389–3402. https://doi.org/10.1093/nar/25.17.3389.

Aravind, L., D. D. Leipe, and E. V. Koonin. 1998. 'Toprim--a Conserved Catalytic Domain in Type IA and II Topoisomerases, DnaG-Type Primases, OLD Family Nucleases and RecR Proteins'. *Nucleic Acids Research* 26 (18): 4205–13. https://doi.org/10.1093/nar/26.18.4205.

Baker, Brett J., Valerie De Anda, Kiley W. Seitz, Nina Dombrowski, Alyson E. Santoro, and Karen G. Lloyd. 2020. 'Diversity, Ecology and Evolution of Archaea'. *Nature Microbiology*, May. https://doi.org/10.1038/s41564-020-0715-z.

Berger, J. M., S. J. Gamblin, S. C. Harrison, and J. C. Wang. 1996. 'Structure and Mechanism of DNA Topoisomerase II'. *Nature* 379 (6562): 225–32. https://doi.org/10.1038/379225a0.

Bush, Natassja G., Katherine Evans-Roberts, and Anthony Maxwell. 2015. 'DNA Topoisomerases'. *EcoSal Plus* 6 (2). https://doi.org/10.1128/ecosalplus.ESP-0010-2014.

Catchpole, Ryan J., and Patrick Forterre. 2019. 'The Evolution of Reverse Gyrase Suggests a Nonhyperthermophilic Last Universal Common Ancestor'. *Molecular Biology and Evolution* 36 (12): 2737–47. https://doi.org/10.1093/molbev/msz180.

Cheung, Kevin J., Vasudeo Badarinarayana, Douglas W. Selinger, Daniel Janse, and George M. Church. 2003. 'A Microarray-Based Antibiotic Screen Identifies a Regulatory Role for Supercoiling in the Osmotic Stress Response of Escherichia Coli'. *Genome Research* 13 (2): 206–15. https://doi.org/10.1101/gr.401003.

Coleman, Gareth A., Adrián A. Davín, Tara A. Mahendrarajah, Lénárd L. Szánthó, Anja Spang, Philip Hugenholtz, Gergely J. Szöllősi, and Tom A. Williams. 2021. 'A Rooted Phylogeny Resolves Early Bacterial Evolution'. *Science (New York, N.Y.)* 372 (6542): eabe0511. https://doi.org/10.1126/science.abe0511.

Corbett, Kevin D., Ryan K. Shultzaberger, and James M. Berger. 2004. 'The C-Terminal Domain of DNA Gyrase A Adopts a DNA-Bending β-Pinwheel Fold'. *Proceedings of the National Academy of Sciences* 101 (19): 7293–98. https://doi.org/10.1073/pnas.0401595101.

Criscuolo, Alexis, and Simonetta Gribaldo. 2010. 'BMGE (Block Mapping and Gathering with Entropy): A New Software for Selection of Phylogenetic Informative Regions from Multiple Sequence Alignments'. *BMC Evolutionary Biology* 10 (July): 210. https://doi.org/10.1186/1471-2148-10-210.

Da Cunha, Violette, Morgan Gaia, Daniele Gadelle, Arshan Nasir, and Patrick Forterre. 2017. 'Lokiarchaea Are Close Relatives of Euryarchaeota, Not Bridging the Gap between Prokaryotes and Eukaryotes'. *PLoS Genetics* 13 (6): e1006810. https://doi.org/10.1371/journal.pgen.1006810.

Dessimoz, Christophe, and Manuel Gil. 2010. 'Phylogenetic Assessment of Alignments Reveals Neglected Tree Signal in Gaps'. *Genome Biology* 11 (4): R37. https://doi.org/10.1186/gb-2010-11-4-r37.

Di Tommaso, Paolo, Sebastien Moretti, Ioannis Xenarios, Miquel Orobitg, Alberto Montanyola, Jia-Ming Chang, Jean-François Taly, and Cedric Notredame. 2011. 'T-Coffee: A Web Server for the Multiple Sequence Alignment of Protein and RNA Sequences Using Structural

Information and Homology Extension'. *Nucleic Acids Research* 39 (Web Server issue): W13-17. https://doi.org/10.1093/nar/gkr245.
Dorman, Charles J. 2006. 'DNA Supercoiling and Bacterial Gene Expression'. *Science Progress* 89 (Pt 3-4): 151–66.
Dorman, Charles J., and Matthew J. Dorman. 2016. 'DNA Supercoiling Is a Fundamental Regulatory Principle in the Control of Bacterial Gene Expression'. *Biophysical Reviews* 8 (3): 209–20. https://doi.org/10.1007/s12551-016-0205-y.
Dress, Andreas W. M., Christoph Flamm, Guido Fritzsch, Stefan Grünewald, Matthias Kruspe, Sonja J. Prohaska, and Peter F. Stadler. 2008. 'Noisy: Identification of Problematic Columns in Multiple Sequence Alignments'. *Algorithms for Molecular Biology: AMB* 3 (June): 7. https://doi.org/10.1186/1748-7188-3-7.
Drlica, K. 1992. 'Control of Bacterial DNA Supercoiling'. *Molecular Microbiology* 6 (4): 425–33. https://doi.org/10.1111/j.1365-2958.1992.tb01486.x.
Forterre, P., G. Mirambeau, C. Jaxel, M. Nadal, and M. Duguet. 1985. 'High Positive Supercoiling in Vitro Catalyzed by an ATP and Polyethylene Glycol-Stimulated Topoisomerase from Sulfolobus Acidocaldarius.' *The EMBO Journal* 4 (8): 2123–28. https://doi.org/10.1002/j.1460-2075.1985.tb03902.x.
Forterre, Patrick. 2002. 'A Hot Story from Comparative Genomics: Reverse Gyrase Is the Only Hyperthermophile-Specific Protein'. *Trends in Genetics: TIG* 18 (5): 236–37.
Forterre, Patrick, Simonetta Gribaldo, Danièle Gadelle, and Marie-Claude Serre. 2007. 'Origin and Evolution of DNA Topoisomerases'. *Biochimie*, DNA Topology, 89 (4): 427–46. https://doi.org/10.1016/j.biochi.2006.12.009.
Gajiwala, Ketan S, and Stephen K Burley. 2000. 'Winged Helix Proteins'. *Current Opinion in Structural Biology* 10 (1): 110–16. https://doi.org/10.1016/S0959-440X(99)00057-3.
Gellert, M., K. Mizuuchi, M. H. O'Dea, and H. A. Nash. 1976. 'DNA Gyrase: An Enzyme That Introduces Superhelical Turns into DNA'. *Proceedings of the National Academy of Sciences of the United States of America* 73 (11): 3872–76. https://doi.org/10.1073/pnas.73.11.3872.
Gilchrist, Cameron L. M., and Yit-Heng Chooi. 2021. 'Clinker & Clustermap.Js: Automatic Generation of Gene Cluster Comparison Figures'. *Bioinformatics (Oxford, England)*, January, btab007. https://doi.org/10.1093/bioinformatics/btab007.
Guindon, Stéphane, Jean-François Dufayard, Vincent Lefort, Maria Anisimova, Wim Hordijk, and Olivier Gascuel. 2010. 'New Algorithms and Methods to Estimate Maximum-Likelihood Phylogenies: Assessing the Performance of PhyML 3.0'. *Systematic Biology* 59 (3): 307–21. https://doi.org/10.1093/sysbio/syq010.
Harrison, S. C., and A. K. Aggarwal. 1990. 'DNA Recognition by Proteins with the Helix-Turn-Helix Motif'. *Annual Review of Biochemistry* 59: 933–69. https://doi.org/10.1146/annurev.bi.59.070190.004441.
Hasegawa, M, and H Kishino. 1994. 'Accuracies of the Simple Methods for Estimating the Bootstrap Probability of a Maximum-Likelihood Tree'. *Molecular Biology and Evolution* 11 (1): 142. https://doi.org/10.1093/oxfordjournals.molbev.a040097.
Hoang, Diep Thi, Olga Chernomor, Arndt von Haeseler, Bui Quang Minh, and Le Sy Vinh. 2018. 'UFBoot2: Improving the Ultrafast Bootstrap Approximation'. *Molecular Biology and Evolution* 35 (2): 518–22. https://doi.org/10.1093/molbev/msx281.
Hocher, Antoine, Maria Rojec, Jacob B. Swadling, Alexander Esin, and Tobias Warnecke. 2019. 'The DNA-Binding Protein HTa from Thermoplasma Acidophilum Is an Archaeal Histone Analog'. *ELife* 8 (November). https://doi.org/10.7554/eLife.52542.
Kalyaanamoorthy, Subha, Bui Quang Minh, Thomas K. F. Wong, Arndt von Haeseler, and Lars S. Jermiin. 2017. 'ModelFinder: Fast Model Selection for Accurate Phylogenetic Estimates'. *Nature Methods* 14 (6): 587–89. https://doi.org/10.1038/nmeth.4285.
Katoh, Kazutaka, and Daron M. Standley. 2013. 'MAFFT Multiple Sequence Alignment Software Version 7: Improvements in Performance and Usability'. *Molecular Biology and Evolution* 30 (4): 772–80. https://doi.org/10.1093/molbev/mst010.

Kikuchi, A., and K. Asai. 1984. 'Reverse Gyrase--a Topoisomerase Which Introduces Positive Superhelical Turns into DNA'. *Nature* 309 (5970): 677–81.
Kishino, H., and M. Hasegawa. 1989. 'Evaluation of the Maximum Likelihood Estimate of the Evolutionary Tree Topologies from DNA Sequence Data, and the Branching Order in Hominoidea'. *Journal of Molecular Evolution* 29 (2): 170–79. https://doi.org/10.1007/BF02100115.
Kishino, Hirohisa, Takashi Miyata, and Masami Hasegawa. 1990. 'Maximum Likelihood Inference of Protein Phylogeny and the Origin of Chloroplasts'. *Journal of Molecular Evolution* 31 (2): 151–60. https://doi.org/10.1007/BF02109483.
Kramlinger, Valerie M., and Hiroshi Hiasa. 2006. 'The "GyrA-Box" Is Required for the Ability of DNA Gyrase to Wrap DNA and Catalyze the Supercoiling Reaction'. *The Journal of Biological Chemistry* 281 (6): 3738–42. https://doi.org/10.1074/jbc.M511160200.
Lal, Avantika, Amlanjyoti Dhar, Andrei Trostel, Fedor Kouzine, Aswin S. N. Seshasayee, and Sankar Adhya. 2016. 'Genome Scale Patterns of Supercoiling in a Bacterial Chromosome'. *Nature Communications* 7 (March). https://doi.org/10.1038/ncomms11055.
Lanz, Martin A., and Dagmar Klostermeier. 2012. 'The GyrA-Box Determines the Geometry of DNA Bound to Gyrase and Couples DNA Binding to the Nucleotide Cycle'. *Nucleic Acids Research* 40 (21): 10893–903. https://doi.org/10.1093/nar/gks852.
Lawrence, J. G., and J. R. Roth. 1996. 'Selfish Operons: Horizontal Transfer May Drive the Evolution of Gene Clusters'. *Genetics* 143 (4): 1843–60. https://doi.org/10.1093/genetics/143.4.1843.
Letunic, Ivica, and Peer Bork. 2021. 'Interactive Tree Of Life (ITOL) v5: An Online Tool for Phylogenetic Tree Display and Annotation'. *Nucleic Acids Research* 49 (W1): W293–96. https://doi.org/10.1093/nar/gkab301.
Liu, L. F., and J. C. Wang. 1987. 'Supercoiling of the DNA Template during Transcription'. *Proceedings of the National Academy of Sciences of the United States of America* 84 (20): 7024–27.
Liu, Yang, Kira S. Makarova, Wen-Cong Huang, Yuri I. Wolf, Anastasia N. Nikolskaya, Xinxu Zhang, Mingwei Cai, et al. 2021. 'Expanded Diversity of Asgard Archaea and Their Relationships with Eukaryotes'. *Nature* 593 (7860): 553–57. https://doi.org/10.1038/s41586-021-03494-3.
López-García, Purificación, Yvan Zivanovic, Philippe Deschamps, and David Moreira. 2015. 'Bacterial Gene Import and Mesophilic Adaptation in Archaea'. *Nature Reviews. Microbiology* 13 (7): 447–56. https://doi.org/10.1038/nrmicro3485.
MacLean, R. Craig, and Alvaro San Millan. 2019. 'The Evolution of Antibiotic Resistance'. *Science (New York, N.Y.)* 365 (6458): 1082–83. https://doi.org/10.1126/science.aax3879.
Martis B, Shiny, Raphaël Forquet, Sylvie Reverchon, William Nasser, and Sam Meyer. 2019. 'DNA Supercoiling: An Ancestral Regulator of Gene Expression in Pathogenic Bacteria?' *Computational and Structural Biotechnology Journal* 17: 1047–55. https://doi.org/10.1016/j.csbj.2019.07.013.
McKay, D. B., and T. A. Steitz. 1981. 'Structure of Catabolite Gene Activator Protein at 2.9 A Resolution Suggests Binding to Left-Handed B-DNA'. *Nature* 290 (5809): 744–49. https://doi.org/10.1038/290744a0.
McKie, Shannon J., Keir C. Neuman, and Anthony Maxwell. 2021. 'DNA Topoisomerases: Advances in Understanding of Cellular Roles and Multi-Protein Complexes via Structure-Function Analysis'. *BioEssays: News and Reviews in Molecular, Cellular and Developmental Biology*, January, e2000286. https://doi.org/10.1002/bies.202000286.
Medini, Duccio, Claudio Donati, Hervé Tettelin, Vega Masignani, and Rino Rappuoli. 2005. 'The Microbial Pan-Genome'. *Current Opinion in Genetics & Development* 15 (6): 589–94. https://doi.org/10.1016/j.gde.2005.09.006.
Mendler, Kerrin, Han Chen, Donovan H Parks, Briallen Lobb, Laura A Hug, and Andrew C Doxey. 2019. 'AnnoTree: Visualization and Exploration of a Functionally Annotated Microbial Tree of Life'. *Nucleic Acids Research* 47 (9): 4442–48. https://doi.org/10.1093/nar/gkz246.

Menzel, Rolf, and Martin Gellert. 1983. 'Regulation of the Genes for E. Coli DNA Gyrase: Homeostatic Control of DNA Supercoiling'. *Cell* 34 (1): 105–13. https://doi.org/10.1016/0092-8674(83)90140-X.

Nguyen, Lam-Tung, Heiko A. Schmidt, Arndt von Haeseler, and Bui Quang Minh. 2015. 'IQ-TREE: A Fast and Effective Stochastic Algorithm for Estimating Maximum-Likelihood Phylogenies'. *Molecular Biology and Evolution* 32 (1): 268–74. https://doi.org/10.1093/molbev/msu300.

Novikova, Olga, Natalya Topilina, and Marlene Belfort. 2014. 'Enigmatic Distribution, Evolution, and Function of Inteins'. *The Journal of Biological Chemistry* 289 (21): 14490–97. https://doi.org/10.1074/jbc.R114.548255.

Oberto, Jacques. 2013. 'SyntTax: A Web Server Linking Synteny to Prokaryotic Taxonomy'. *BMC Bioinformatics* 14 (January): 4. https://doi.org/10.1186/1471-2105-14-4.

Parks, Donovan H., Maria Chuvochina, Pierre-Alain Chaumeil, Christian Rinke, Aaron J. Mussig, and Philip Hugenholtz. 2020. 'A Complete Domain-to-Species Taxonomy for Bacteria and Archaea'. *Nature Biotechnology*, April, 1–8. https://doi.org/10.1038/s41587-020-0501-8.

Peter, Brian J., Javier Arsuaga, Adam M. Breier, Arkady B. Khodursky, Patrick O. Brown, and Nicholas R. Cozzarelli. 2004. 'Genomic Transcriptional Response to Loss of Chromosomal Supercoiling in Escherichia Coli'. *Genome Biology* 5 (11): R87. https://doi.org/10.1186/gb-2004-5-11-r87.

Petitjean, Céline, Philippe Deschamps, Purificación López-García, and David Moreira. 2014. 'Rooting the Domain Archaea by Phylogenomic Analysis Supports the Foundation of the New Kingdom Proteoarchaeota'. *Genome Biology and Evolution* 7 (1): 191–204. https://doi.org/10.1093/gbe/evu274.

Polz, Martin F., Eric J. Alm, and William P. Hanage. 2013. 'Horizontal Gene Transfer and the Evolution of Bacterial and Archaeal Population Structure'. *Trends in Genetics: TIG* 29 (3): 170–75. https://doi.org/10.1016/j.tig.2012.12.006.

Pommier, Yves, ed. 2012. *DNA Topoisomerases and Cancer*. Cancer Drug Discovery and Development. Humana Press. //www.springer.com/it/book/9781461403227.

Pruss, G. J., S. H. Manes, and K. Drlica. 1982. 'Escherichia Coli DNA Topoisomerase I Mutants: Increased Supercoiling Is Corrected by Mutations near Gyrase Genes'. *Cell* 31 (1): 35–42. https://doi.org/10.1016/0092-8674(82)90402-0.

Qi, Yuan, Jimin Pei, and Nick V. Grishin. 2002. 'C-Terminal Domain of Gyrase A Is Predicted to Have a Beta-Propeller Structure'. *Proteins* 47 (3): 258–64. https://doi.org/10.1002/prot.10090.

Raymann, Kasie, Céline Brochier-Armanet, and Simonetta Gribaldo. 2015. 'The Two-Domain Tree of Life Is Linked to a New Root for the Archaea'. *Proceedings of the National Academy of Sciences of the United States of America* 112 (21): 6670–75. https://doi.org/10.1073/pnas.1420858112.

Raymann, Kasie, Patrick Forterre, Céline Brochier-Armanet, and Simonetta Gribaldo. 2014. 'Global Phylogenomic Analysis Disentangles the Complex Evolutionary History of DNA Replication in Archaea'. *Genome Biology and Evolution* 6 (1): 192–212. https://doi.org/10.1093/gbe/evu004.

Reece, R. J., and A. Maxwell. 1991. 'DNA Gyrase: Structure and Function'. *Critical Reviews in Biochemistry and Molecular Biology* 26 (3–4): 335–75. https://doi.org/10.3109/10409239109114072.

Robert, Xavier, and Patrice Gouet. 2014. 'Deciphering Key Features in Protein Structures with the New ENDscript Server'. *Nucleic Acids Research* 42 (Web Server issue): W320-324. https://doi.org/10.1093/nar/gku316.

Rodríguez-Beltrán, Jerónimo, Vidar Sørum, Macarena Toll-Riera, Carmen de la Vega, Rafael Peña-Miller, and Álvaro San Millán. 2020. 'Genetic Dominance Governs the Evolution and Spread of Mobile Genetic Elements in Bacteria'. *Proceedings of the National Academy of Sciences* 117 (27): 15755–62. https://doi.org/10.1073/pnas.2001240117.

Rui, Shan, and Yuk-Ching Tse-Dinh. 2003. 'Topoisomerase Function during Bacterial Responses to Environmental Challenge'. *Frontiers in Bioscience: A Journal and Virtual Library* 8 (January): d256-263. https://doi.org/10.2741/984.

Schoeffler, Allyn J., and James M. Berger. 2008. 'DNA Topoisomerases: Harnessing and Constraining Energy to Govern Chromosome Topology'. *Quarterly Reviews of Biophysics* 41 (1): 41–101. https://doi.org/10.1017/S003358350800468X.

Seol, Yeonee, and Keir C. Neuman. 2016. 'The Dynamic Interplay Between DNA Topoisomerases and DNA Topology'. *Biophysical Reviews* 8 (3): 221–31. https://doi.org/10.1007/s12551-016-0206-x.

Shimodaira, H, and M Hasegawa. 1999. 'Multiple Comparisons of Log-Likelihoods with Applications to Phylogenetic Inference'. *Molecular Biology and Evolution* 16 (8): 1114. https://doi.org/10.1093/oxfordjournals.molbev.a026201.

Shimodaira, Hidetoshi. 2002. 'An Approximately Unbiased Test of Phylogenetic Tree Selection'. *Systematic Biology* 51 (3): 492–508. https://doi.org/10.1080/10635150290069913.

Sioud, M, G. Baldacci, A.M. de Recondo, and P. Forterre. 1988. 'Novobiocin Induces Positive Supercoilling of Small Plasmids from Halophilic Arctiaebacterla in Vivo'. *Nucleic Acids Research* 16 (4): 1379–91. https://doi.org/10.1093/nar/16.4.1379.

Sioud, M, O Possot, C Elie, L Sibold, and P Forterre. 1988. 'Coumarin and Quinolone Action in Archaebacteria: Evidence for the Presence of a DNA Gyrase-like Enzyme.' *Journal of Bacteriology* 170 (2): 946–53. https://doi.org/10.1128/JB.170.2.946-953.1988.

Soczek, Katarzyna M., Tim Grant, Peter B. Rosenthal, and Alfonso Mondragón. 2018. 'CryoEM Structures of Open Dimers of Gyrase A in Complex with DNA Illuminate Mechanism of Strand Passage'. *ELife* 7 (November): e41215. https://doi.org/10.7554/eLife.41215.

Steinegger, Martin, and Johannes Söding. 2017. 'MMseqs2 Enables Sensitive Protein Sequence Searching for the Analysis of Massive Data Sets'. *Nature Biotechnology* 35 (11): 1026–28. https://doi.org/10.1038/nbt.3988.

Straney, R., R. Krah, and R. Menzel. 1994. 'Mutations in the -10 TATAAT Sequence of the GyrA Promoter Affect Both Promoter Strength and Sensitivity to DNA Supercoiling'. *Journal of Bacteriology* 176 (19): 5999–6006. https://doi.org/10.1128/jb.176.19.5999-6006.1994.

Strimmer, Korbinian, and Andrew Rambaut. 2002. 'Inferring Confidence Sets of Possibly Misspecified Gene Trees'. *Proceedings. Biological Sciences* 269 (1487): 137–42. https://doi.org/10.1098/rspb.2001.1862.

Sutormin, Dmitry, Natalia Rubanova, Maria Logacheva, Dmitry Ghilarov, and Konstantin Severinov. 2018. 'Single-Nucleotide-Resolution Mapping of DNA Gyrase Cleavage Sites across the Escherichia Coli Genome'. *Nucleic Acids Research*, December. https://doi.org/10.1093/nar/gky1222.

Takahashi, Tomio S., Violette Da Cunha, Mart Krupovic, Claudine Mayer, Patrick Forterre, and Danièle Gadelle. 2020. 'Expanding the Type IIB DNA Topoisomerase Family: Identification of New Topoisomerase and Topoisomerase-like Proteins in Mobile Genetic Elements'. *NAR Genomics and Bioinformatics* 2 (1): lqz021. https://doi.org/10.1093/nargab/lqz021.

Tan, Ge, Matthieu Muffato, Christian Ledergerber, Javier Herrero, Nick Goldman, Manuel Gil, and Christophe Dessimoz. 2015. 'Current Methods for Automated Filtering of Multiple Sequence Alignments Frequently Worsen Single-Gene Phylogenetic Inference'. *Systematic Biology* 64 (5): 778–91. https://doi.org/10.1093/sysbio/syv033.

Tarasov, Valery, Rita Schwaiger, Katarina Furtwängler, Mike Dyall-Smith, and Dieter Oesterhelt. 2011. 'A Small Basic Protein from the Brz-Brb Operon Is Involved in Regulation of Bop Transcription in Halobacterium Salinarum'. *BMC Molecular Biology* 12 (September): 42. https://doi.org/10.1186/1471-2199-12-42.

Tisza, Michael J., and Christopher B. Buck. 2021. 'A Catalog of Tens of Thousands of Viruses from Human Metagenomes Reveals Hidden Associations with Chronic Diseases'. *Proceedings of the National Academy of Sciences of the United States of America* 118 (23): e2023202118. https://doi.org/10.1073/pnas.2023202118.

Tretter, Elsa M., Jeffrey C. Lerman, and James M. Berger. 2010. 'A Naturally Chimeric Type IIA Topoisomerase in Aquifex Aeolicus Highlights an Evolutionary Path for the Emergence of

Functional Paralogs'. *Proceedings of the National Academy of Sciences of the United States of America* 107 (51): 22055–59. https://doi.org/10.1073/pnas.1012938107.
Unniraman, S., and V. Nagaraja. 1999. 'Regulation of DNA Gyrase Operon in Mycobacterium Smegmatis: A Distinct Mechanism of Relaxation Stimulated Transcription'. *Genes to Cells: Devoted to Molecular & Cellular Mechanisms* 4 (12): 697–706. https://doi.org/10.1046/j.1365-2443.1999.00296.x.
Vallenet, David, Alexandra Calteau, Mathieu Dubois, Paul Amours, Adelme Bazin, Mylène Beuvin, Laura Burlot, et al. 2020. 'MicroScope: An Integrated Platform for the Annotation and Exploration of Microbial Gene Functions through Genomic, Pangenomic and Metabolic Comparative Analysis'. *Nucleic Acids Research* 48 (D1): D579–89. https://doi.org/10.1093/nar/gkz926.
Vanden Broeck, Arnaud, Christophe Lotz, Julio Ortiz, and Valérie Lamour. 2019. 'Cryo-EM Structure of the Complete E. Coli DNA Gyrase Nucleoprotein Complex'. *Nature Communications* 10 (1): 4935. https://doi.org/10.1038/s41467-019-12914-y.
Villain, Paul, Violette Da Cunha, Etienne Villain, Patrick Forterre, Jacques Oberto, Ryan Catchpole, and Tamara Basta. 2021. 'The Hyperthermophilic Archaeon Thermococcus Kodakarensis Is Resistant to Pervasive Negative Supercoiling Activity of DNA Gyrase', 16.
Wang, James C. 2002. 'Cellular Roles of DNA Topoisomerases: A Molecular Perspective'. *Nature Reviews. Molecular Cell Biology* 3 (6): 430–40. https://doi.org/10.1038/nrm831.
Ward, Doyle, and Austin Newton. 1997. 'Requirement of Topoisomerase IV *ParC* and *ParE* Genes for Cell Cycle Progression and Developmental Regulation in *Caulobacter Crescentus*'. *Molecular Microbiology* 26 (5): 897–910. https://doi.org/10.1046/j.1365-2958.1997.6242005.x.
Williams, Tom A., Gergely J. Szöllősi, Anja Spang, Peter G. Foster, Sarah E. Heaps, Bastien Boussau, Thijs J. G. Ettema, and T. Martin Embley. 2017. 'Integrative Modeling of Gene and Genome Evolution Roots the Archaeal Tree of Life'. *Proceedings of the National Academy of Sciences of the United States of America* 114 (23): E4602–11. https://doi.org/10.1073/pnas.1618463114.
Yamashiro, Kan, and Akihiko Yamagishi. 2005. 'Characterization of the DNA Gyrase from the Thermoacidophilic Archaeon Thermoplasma Acidophilum'. *Journal of Bacteriology* 187 (24): 8531–36. https://doi.org/10.1128/JB.187.24.8531-8536.2005.

Article II

Nucleic Acids Research, 2021 **1–16**
https://doi.org/10.1093/nar/gkab869

The hyperthermophilic archaeon *Thermococcus kodakarensis* is resistant to pervasive negative supercoiling activity of DNA gyrase

Paul Villain[1], Violette da Cunha[1], Etienne Villain[3], Patrick Forterre[1,2], Jacques Oberto[1], Ryan Catchpole[1,4,*] and Tamara Basta[1,*]

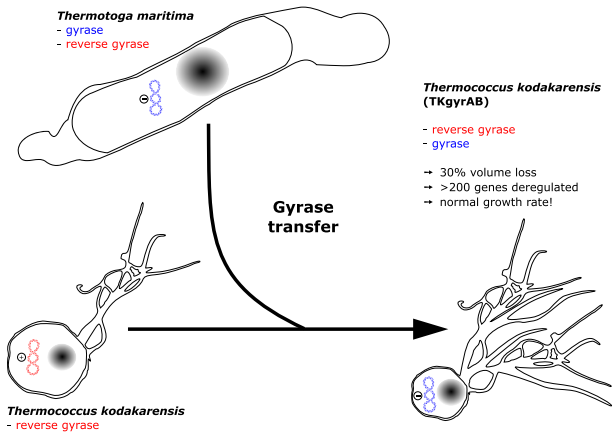
[1]Université Paris-Saclay, CEA, CNRS, Institute for Integrative Biology of the Cell (I2BC), 91198 Gif-sur-Yvette, France, [2]Archaeal Virology Unit, Institut Pasteur, Paris, France, [3]Immunology department, Institut Pasteur, Paris, France and [4]Department of Biochemistry and Molecular Biology, University of Georgia, Athens, GA 30602, USA

Received August 01, 2021; Revised September 10, 2021; Editorial Decision September 14, 2021; Accepted September 30, 2021

ABSTRACT

In all cells, DNA topoisomerases dynamically regulate DNA supercoiling allowing essential DNA processes such as transcription and replication to occur. How this complex system emerged in the course of evolution is poorly understood. Intriguingly, a single horizontal gene transfer event led to the successful establishment of bacterial gyrase in Archaea, but its emergent function remains a mystery. To better understand the challenges associated with the establishment of pervasive negative supercoiling activity, we expressed the gyrase of the bacterium *Thermotoga maritima* in a naïve archaeon *Thermococcus kodakarensis* which naturally has positively supercoiled DNA. We found that the gyrase was catalytically active in *T. kodakarensis* leading to strong negative supercoiling of plasmid DNA which was stably maintained over at least eighty generations. An increased sensitivity of gyrase-expressing *T. kodakarensis* to ciprofloxacin suggested that gyrase also modulated chromosomal topology. Accordingly, global transcriptome analyses revealed large scale gene expression deregulation and identified a subset of genes responding to the negative supercoiling activity of gyrase. Surprisingly, the artificially introduced dominant negative supercoiling activity did not have a measurable effect on *T. kodakarensis* growth rate. Our data suggest that gyrase can become established in *Thermococcales* archaea without critically interfering with DNA transaction processes.

GRAPHICAL ABSTRACT



INTRODUCTION

With the discovery of the structure of DNA, it became apparent that the opening of the double helix generates torsional stress resulting in overwinding or underwinding of the DNA molecule (1). Paradoxically, many DNA transaction processes such as transcription and replication require strand separation and will lead naturally to DNA overwinding and strand entanglement (2–4). These topological constraints antagonise these essential cellular processes and if not resolved, are lethal. To deal with this problem, all cells rely on topoisomerases, a ubiquitous class of enzymes that introduce strand breaks to relieve unfavourable topological intermediates without damaging the genome (5–11). Topoisomerases are mechanistically classified as type I or type II, depending on whether they cleave one or two strands of DNA, respectively (6). Multiple phylogenetically unrelated subclasses of each type exist in the biosphere and such diversity has made it particularly challenging to dissect the evolutionary history of topoisomerases (12,13). A long-standing

*To whom correspondence should be addressed. Tel: +33 1 69 82 61 67; Email: tamara.basta@i2bc.paris-saclay.fr
Correspondence may also be addressed to Ryan Catchpole. Email: ryan.catchpole@uga.edu

puzzle has been to understand why so many topoisomerases have emerged in the course of evolution and what role they played in the evolution of DNA-based cells (12,13).

DNA gyrase (hereafter gyrase), a type II A topoisomerase, is the only known enzyme that can negatively supercoil (underwind) DNA using the free energy of ATP hydrolysis to drive the process (14). An important antibiotic target, gyrase is essential and ubiquitous in bacteria where it controls (together with Topo I) the supercoiling density of chromosomes by introducing negative supercoils into DNA and by relaxing positive supercoils accumulating in front of moving DNA and RNA polymerases (15–18). The contribution of DNA gyrase in maintaining the chromosome in an underwound state in bacterial cells can profoundly impact the binding of regulatory proteins, promoter firing dynamics, DNA replication, and chromosome architecture (4,19,20).

DNA supercoiling is used in a wide range of bacteria to quickly transduce environmental signals towards the chromosome and this process is conserved in distant bacterial species (20,21). The most clearly described pathway involves the modulation of gyrase activity in response to [ATP]/[ADP] ratio in the cell. When this ratio is low, the DNA gyrase supercoiling activity is significantly reduced and the expression level of many genes is simultaneously modified (22,23). The inhibition of gyrase supercoiling activity by quinolone antibiotics has a similar effect; the expression of up to 48% of genes can be deregulated (21). This relatively simple, quick and general mechanism was suggested to be one of the key evolutionary inventions allowing the bacteria to occupy a wide variety of environments (12,13).

In Archaea, Gyrase is found in all members of the highly diversified monophyletic group (named Cluster II Euryarchaeota by Adam and colleagues) containing seven distinct groups with very different lifestyles (acidophiles, halophiles, methanogens among others) and, sporadically, in DPANN and Asgard superphyla (12,24,25) (Figure 1). Initial phylogenetic analyses indicated that archaeal gyrase is of bacterial origin and was acquired via ancient horizontal gene transfer by a hyperthermophilic archaeon (12). Later analysis including more archaeal lineages suggested that this transfer occurred only once at the base of the aforementioned late emerging Cluster II Euryarchaeota (26). Because negative supercoiling facilitates DNA melting, it was proposed that gyrase acquisition had a profound impact on all DNA-dependent processes with important consequences for the evolution of recipient archaea (13,27,28). However, how and why DNA gyrase became fixed in archaeal lineages remains obscure.

The successful establishment of bacterial gyrase in Archaea is particularly intriguing since the introduction of an uncontrolled negative supercoiling activity could potentially interfere with all DNA-templated processes. To make things worse, the recipient archaeon was probably a thermophile (28–31) and therefore encoded a reverse gyrase with opposite, positive supercoiling activity which is essential for life at high temperature (32–35). Finally, archaea encode Topo VI as the main type II topoisomerase and its predicted *in vivo* role in relaxing the positive supercoils overlaps with that of gyrase (36).

To understand better the challenges imposed by DNA gyrase to a naïve archaeal cell we introduced a bacterial gyrase in the genetically tractable hyperthermophilic archaeon *Thermococcus kodakarensis* TS559 which naturally has slightly positively supercoiled DNA (37). This archaeon encodes histones and three topoisomerases (reverse gyrase, Topo III and Topo VI) thus mimicking the topological 'kit' present in the ancestor of the Cluster II Euryarchaeota (13,29). As a source of gyrase, we selected the one from the bacterium *Thermotoga maritima* (TmGyrAB) since its closest relatives are archaeal gyrases (Villain *et al.*, unpublished) increasing the chance that its activity would not be impaired by the archaeal cellular context. In addition, the optimal growth temperature of *T. maritima* (80°C) and that of *T. kodakarensis* (85°C) are similar, both were isolated from geothermally heated sea floors (38,39) and they can be co-cultured in laboratory. Finally, TmGyrAB exhibits the expected negative supercoiling activity both *in vitro* and *in vivo* (40–42).

We found that the gyrase of *T. maritima* was active and predominant in *T. kodakarensis* such that the normally positively supercoiled plasmid DNA was converted to strongly negatively supercoiled DNA. Gyrase interacted with the genome of *T. kodakarensis* and induced differential expression of hundreds of genes, a subset of which specifically responded to negative supercoiling activity. Reverse gyrase was the only topoisomerase that reacted transcriptionally, albeit modestly, to negative supercoiling activity. Despite unnatural gyrase-enforced topological changes, the growth of *T. kodakarensis* was not affected. We conclude that gyrase is remarkably well tolerated by *T. kodakarensis* suggesting the existence of resilience mechanisms against torsional stress which may have been instrumental for the natural establishment of gyrase in the archaeal domain.

MATERIALS AND METHODS

Construction of recombinant *Thermococcus kodakarensis* TS559 strains

All the strains, plasmids and oligonucleotides used in this study are listed in the Supplementary table S1 and S2. Plasmids were constructed in *Escherichia coli* strain XL1-Blue grown at 37°C in LB supplemented with Ampicillin (100 μg/ml), Kanamycin (40 μg/ml) or Chloramphenicol (20 μg/ml) using standard molecular biology protocols. The gyrase-encoding genes *gyrA* and *gyrB* were PCR amplified using genomic DNA of *Thermotoga maritima* MSB8 as template and cloned as a bi-cistronic operon in plasmid pTNAg (43) under the control of the strong constitutive promoter PhmtB. Plasmid constructions were performed by Gibson Assembly using the NEBuilder HiFi DNA Assembly Master Mix (New England Biolabs) following the manufacturer's protocol.

T. kodakarensis TS559 was transformed using standard protocol in anaerobic conditions using the rich medium ASW-YT or the synthetic medium without tryptophan ASW-AAW (43,44). Briefly, 10 ml of late exponential phase culture was pelleted and resuspended in 100 μl of 0.8× ASW. At least 2 μg of plasmid DNA was mixed with 100 μl of cell suspension and incubated on ice for 1 h, heat shocked for 1 min at 85°C and cooled on ice for 10 min.

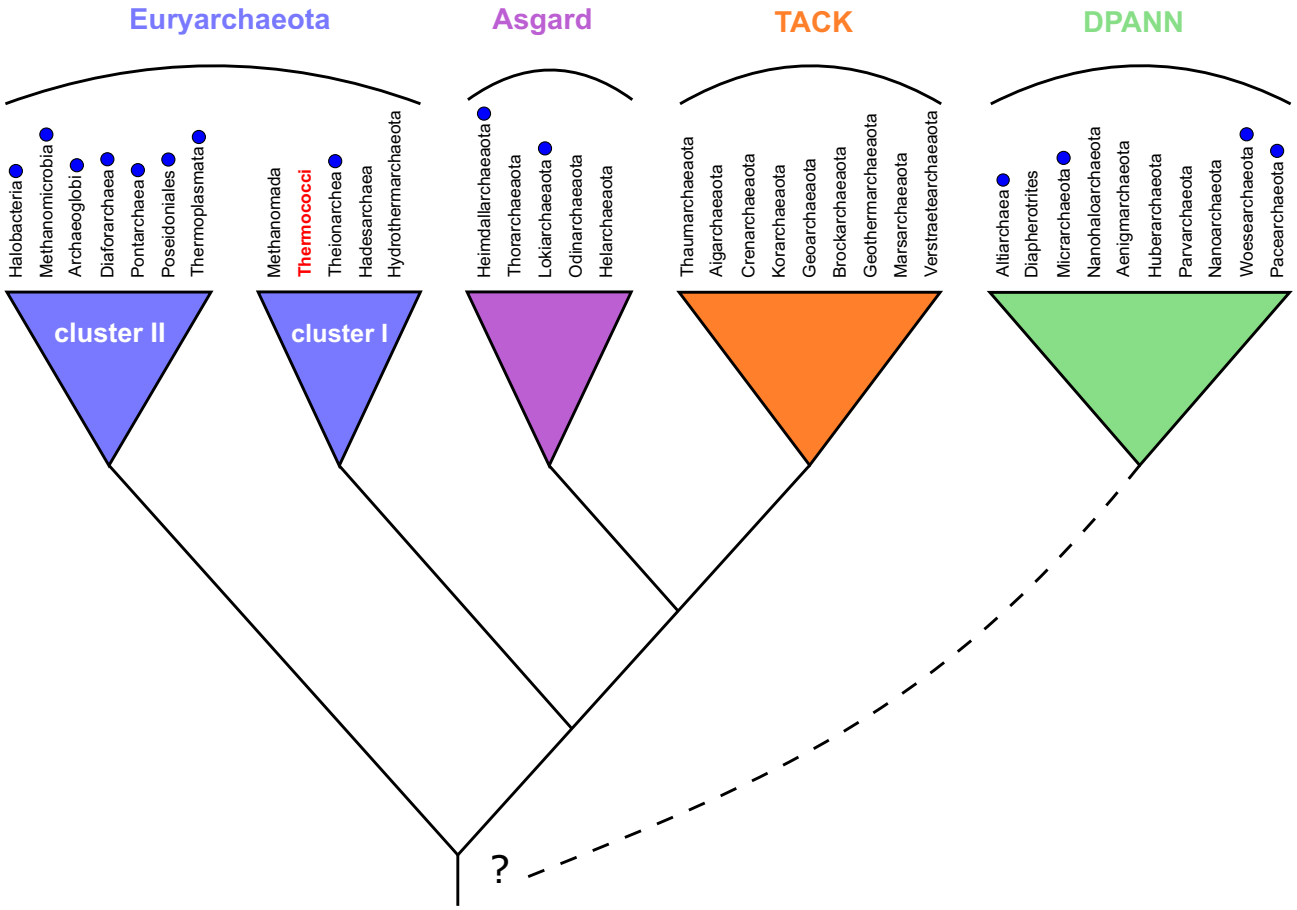


Figure 1. DNA gyrase distribution in Archaea. Schematic phylogeny of archaea with main phyla and superphyla indicated at the top. The presence of gyrase in a phylum is indicated by a blue dot. The model organism in this study, *Thermococcus kodakarenis* KOD1, belongs to Thermococci phylum within Euryarchaeota I clade. The dashed line symbolizes the uncertainty of DPANN branching within the Archaea tree.

Then 1 ml of non-selective medium was added and cell suspensions were incubated 90 min at 85°C for recovery. Cells were then harvested by centrifuging 4 min at 4500 g, resuspended in 200 µl of 0.8× ASW and plated onto selective solid medium containing 1% (w/v) Phytagel™ (Sigma-Aldrich). Plates were reduced with 2 ml of polysulfide solution (100 g of Na_2S nonahydrate and 30 g of sulfur dissolved in 75 ml of water) per litre of medium and starch azure was added at 0.2% (w/v) to facilitate visualisation of colonies (45). After 40 h (ASW-YT) or 64 h (ASW-AAW$^-$) of incubation at 85°C isolated colonies were transferred to selective liquid medium in sealed bottles under N_2 atmosphere. Na_2S was added to a final concentration of 0.02% (w/v) and resazurin was added at 1 mg/l as an indicator of medium reduction.

To measure growth, liquid cultures of *T. kodakarensis* were inoculated at 1:100 dilution from fresh precultures. Cell density was monitored with phase-contrast microscopy using a Thoma cell counting chamber (0.01 mm depth) or using an optical device that measured turbidity variations directly in Hungate tubes (MicrobeMeter).

The expression of the *gyrA* and *gyrB* genes in transformants was tested by RT-PCR. Briefly, three individual clones of *T. kodakarensis* TS559 were each grown in 25 ml of ASW-YT without agmatine until late exponential phase. Cells were harvested (5000 × g, 10 min) and resuspended in 500 µl of resuspension buffer (NaCl 1 M, Tris HCl 0.1 M, $CaCl_2$ 5 mM, $MgSO_4$ 0.1 M). The RNAs were extracted using TRIzol according to the supplied protocol (Sigma-Aldrich). To eliminate residual DNA, the samples were treated for 30 min at 37°C with TURBO DNase (ThermoFisher) and reextracted with TRIzol. The cDNA was synthetized using Maxima First Strand cDNA Synthesis kit (ThermoFisher) following the manufacturer's protocol. The PCR reactions were performed using each cDNA or total RNA sample as template. The used oligonucleotides are listed in Supplementary table S2. The obtained PCR products were separated on 1% (w/v) agarose gel and stained with ethidium bromide (0.5 µg/ml).

Wide-field microscopy and DNA staining of *T. kodakarensis* cells

A small volume (200 µl) of exponentially growing culture was rapidly cooled down in mixture of ice and water to limit the effects of cold shock and oxygen exposure during imaging. Cells were then centrifuged and resuspended in an equal volume of 0.8× ASW containing 1 µg/ml of Hoechst 33342 (excitation 361 nm / emission 486 nm) for DNA staining. Cells were incubated for 10 min in the dark

and then mounted onto glass slides covered with a thin layer of 1% agarose (suspended in 0.8× ASW solution). DIC and fluorescent images were obtained at room temperature on an SP8 confocal laser scanning microscope (Leica Microsystems) equipped with hybrid detectors and a 63x oil immersion objective (HC Plan Apo, 1.4 numerical aperture [NA]; Leica). Fluorescence detection was performed by exciting the sample with a 405 nm laser and collecting fluorescence between 415 and 515 nm) at the speed of 600 Hz with a line averaging of three. Image format was adjusted to provide an XY optimal sampling (pixel size of 60 nm) and for each position z-stacks (3 μm width and 0.5 μm step) were acquired.

Method Yen from MicrobeJ software (46) was used to perform automatic cell detection and size measurements. The obtained profiles were manually curated to remove debris and aggregated cells. For cell size, circularity and area measurements at least 50 independent images were analysed per strains totalling at least 600 analysed cells.

Two-dimensional agarose gel electrophoresis

Plasmids pTNAg-Y119F and pTNAg-gyrAB are too large (12 602 bp) to be easily resolved by 2D agarose gel electrophoresis we instead used the smaller reporter plasmid pTPTK2 (5455 bp) (43). *T. kodakarensis* strains were grown in Ravot medium (47) to increase the yield of intact supercoiled plasmids.

Ravot medium was inoculated from overnight precultures grown in ASW-AAW medium at 1:100 dilution, and growth was monitored until late exponential phase. Then cultures were rapidly chilled in precooled beaker immersed in a water-ice bath to stop topoisomerase or nuclease activities. Plasmids were extracted with the NucleoSpin® kit (Macherey-Nagel) following the low copy manufacturer's protocol with minor modifications: (i) the lysis step was reduced to 1 min to limit plasmid nicking, (ii) two steps of lysate clarification were performed and (iii) the wash step (AW) was performed to inactivate residual nucleases.

Agarose gels were prepared by dissolving 0.8% (w/v) of ultrapure agarose (Invitrogen) in 1× TBE buffer (89 mM Tris, 89 mM boric acid, 2 mM EDTA) and pouring in 14.5 × 20 cm tray. Electrophoresis was performed at 24°C and minimum 3 μg of plasmid DNA was used for each sample. In the first dimension, no intercalating agent or 1.5 μg/ml of chloroquine (Alpha Aesar) was added in the gel and in the running buffer. Electrophoresis was run at 1.2 V/cm for 40 h. Gels were subsequently equilibrated for 1 h in 1× TBE buffer supplemented with 7.5 μg/ml of chloroquine and placed in the tray after a rotation of 90°. In the second dimension, an electric field of 2 V/cm was applied for 10 h. Gels were washed 3 × 30 min in water to remove chloroquine and then stained for 1 h with 2.5 μg/ml of ethidium bromide. Gels were then rinsed with water and imaged with a Typhoon Imager (Amersham) using Cy3 channel.

Calculation of the supercoiling density

Superhelical density (σ) of the reporter plasmid pTPTK2 (5455 bp) was calculated from the imaged 2D gels using an adaptation of the band-counting method as described by López-García and Forterre (48). As an example, we detail the calculation for the gels presented in Figure 2. Topoisomers were first separated without chloroquine in the first dimension and the major topoisomer was identified using band intensity measurement with Fiji software (49). This gave the writhe for the plasmid pTPTK2 (Wr = –5) in strain TKY119F. A second gel was run with 1.5 μg/ml of chloroquine in the first dimension, allowing to determine a Wr = + 14 for plasmid pTPTK2 in strain TKY119F. We thus determined that the chloroquine introduced 19 positive supercoils in pTPTK2. To account for the temperature effect on plasmid topology between the growth temperature of *Thermococcus* (85°C) and the electrophoresis (24°C) we applied the correcting factor of - 0.011°/°C/bp (50). The temperature difference thus introduced 10 negative supercoils in pTPTK2. Taking into account the chloroquine effect in the first dimension and the unwinding induced by temperature, we determined a Wr of –5 for the TKY119F strain (Wr = 14 + 10 – 19) and -17 for the TKgyrAB strain (–8 + 10 – 19). Tw is equal to the length of the plasmid molecule (in base pairs) divided by the number of base pairs per helix turn (h) (h = 10.5 bp per turn under standard conditions). For pTPTK2 the twist is Tw = 5455 bp/10.5 bp. Finally, the σ was calculated by dividing the Wr by Tw to yield $\sigma_{TKY119F}$ = +0.0096 and $\sigma_{TKgyrAB}$ = –0.0327. The mean σ values were calculated from three independent 2D gel electrophoresis experiments.

Plasmid toxicity assay

T. kodakarensis strains TKY119F and TKgyrAB, carrying respectively pLCAg-Y119F and pLCAg-gyrAB, were passaged 14 times at 85°C with 100-fold dilution in Ravot medium or in Ravot medium supplemented with 1 mM of agmatine. The culture used for the first inoculation and the 4th, 9th and 14th subcultures, corresponding respectively to 0, 24, 54 and 84 cell generations, were anaerobically sampled. These samples were 10-fold serial diluted in 1X Ravot salts solution to a 10^{-7} dilution. 10 μl of each dilution were spotted on solid selective (Ravot) or non-selective (Ravot with agmatine) medium. Plates were reduced with polysulfides solution and supplemented with starch azure as described above. After 2 days of incubation in anaerobic jars at 85°C, cell viability was determined by counting colonies from spotted dilutions.

Ciprofloxacin susceptibility test

Ciprofloxacin sensitivity of TKY119F and TKgyrAB was investigated on plates using an adaptation of the inverted spot test method (47). Briefly, in anaerobic conditions, 25 ml of Ravot-phytagel medium supplemented with agmatine was poured in Petri dishes as described above. On this bottom layer, a 5 ml thin layer composed of 1X Ravot medium, 0.18% (w/v) of Phytagel™, 1 mM of agmatine, 1% of colloidal sulfur and 160 μl of late exponential phase *T. kodakarensis* culture was poured. On the solidified top layer, 5 mm Whatman paper discs were gently laid. 10 μl of ciprofloxacin dilutions dissolved in 0.1 N HCl were spotted on the paper discs. On such plates the top layer initially appears milky-white because of the colloidal sulfur. Growing

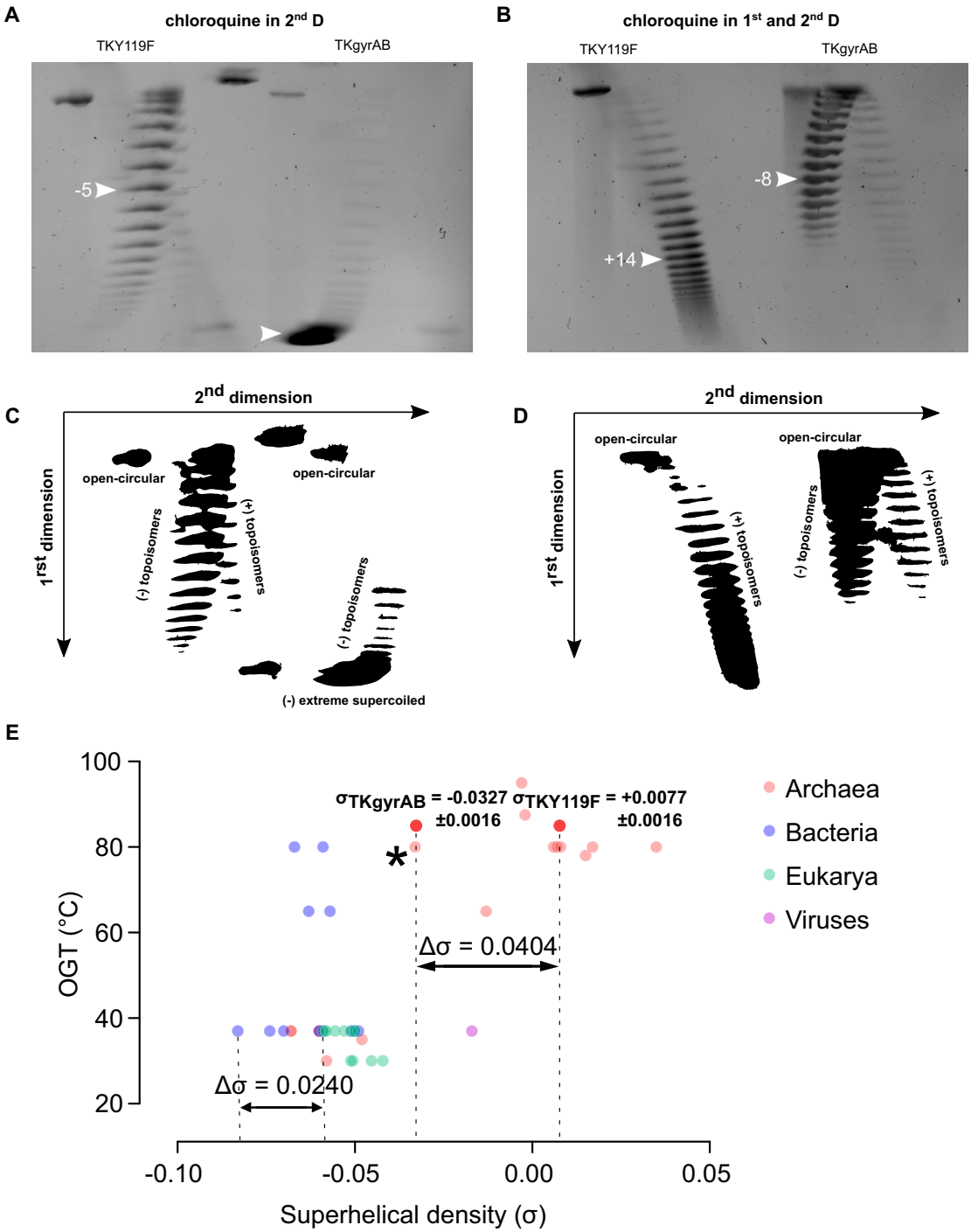


Figure 2. Plasmid DNA from *Thermococcus kodakarensis* TKgyrAB is negatively supercoiled. Reporter plasmid pTPTK2 was isolated from TKY119F or TKgyrAB strains and its topoisomers were separated using 2D agarose gel electrophoresis. The corresponding cartoon is depicted below each agarose gel. (**A**) and (**C**) DNA intercalating drug chloroquine was added at 7.5 μg/ml only in the second-dimension run. This allows to separate positively and negatively supercoiled topoisomers as fast-migrating right arc and slow-migrating left arc, respectively. Under these running conditions, the majority of pTPTK2 isolated from TKgyrAB strain is in extremely negatively supercoiled form. The major (the most abundant) topoisomer from the TKY119F is indicated with a white arrow and was used to calculate superhelical density. (**B** and **D**) Chloroquine was added both in the first (1.5 μg/ml) and the second-dimension run (7.5 μg/ml). Chloroquine introduces a slight positive torsion in DNA, thus changing the apparent superhelical density of topoisomers toward a positively supercoiled state. This allowed relaxing the extremely negatively supercoiled form of pTPTK2 during the first dimension run such that the individual topoisomers could be separated and the major (the most abundant) topoisomer determined. This topoisomer is indicated by a white arrow and was used to calculate superhelical density. (**E**) Plasmid superhelical densities from various organisms were plotted against the optimal growth temperature of their hosts (see also supplementary table S3). The mean σ and standard deviation calculated from three independent experiments are indicated for gyrase expressing strain TKgyrAB and the control strains TKY119F. The point corresponding to Archaeoglobi archaea is highlighted by an asterisk. The total change in the supercoiling density (Δσ) resulting from gyrase activity in *T. kodakarensis* or from TopoI inhibition in *Streptococcus pneumonia* is indicated.

T. kodakarensis consumes the sulfur and the top layer becomes transparent. After overnight incubation in anaerobic jars, growth inhibition was assessed by measuring the diameter of the remaining milky-white halos around the Whatman paper discs.

Differential gene expression analysis

To prepare samples for RNA sequencing, 25 ml of Ravot medium was inoculated at 1/100 dilution with fresh *T. kodakarensis* preculture. Total RNA was extracted from 20 ml of exponentially growing cultures (6 h of culture, approximately 2×10^9 cells/ml) using a NucleoSpin RNA set for NucleoZOL (Macherey Nagel). DNA was eliminated from samples using a TURBO DNA-free kit following the manufacturers protocol (Ambion). For each strain, biological replicates were prepared from four independent cultures. Total RNA quality was assessed on an Agilent Bioanalyzer 2100, using RNA 6000 pico kit (Agilent Technologies). Directional RNA-Seq Libraries were constructed using the TruSeq Stranded Total RNA library prep kit, with bacteria Ribo-Zero reagents (Illumina), following the manufacturer's instructions; 500 ng of total RNA were used. After the Ribo-Zero step, the samples were checked on the Agilent Bioanalyzer for proper rRNA depletion. Final libraries quality was assessed on an Agilent Bioanalyzer 2100, using an Agilent High Sensitivity DNA Kit. Libraries were pooled in equimolar proportions and sequenced on a Paired-End 2 × 75 bp run, on an Illumina NextSeq500 instrument. Demultiplexing was performed with bcl2fastq2 v2.18.12. Adapters were trimmed with Cutadapt v1.15, and only reads longer than 10 bp were kept for further analysis. Between 32 and 56 million reads for each sample were mapped on genome of *T. kodakarensis* TS559 and plasmid sequences with Burrow-Wheeler Aligner for short-read alignment release 0.7.17-r1188 (51). Reads per gene were counted using subread featureCounts v1.5.2 and differential analyses were performed in R using DESeq2 v1.28.1 package (52).

Adjusted *P*-values histograms, Principal Component Analysis, Volcano Plot, Heatmap and distribution of deregulated genes across *T. kodakarensis* TS559 chromosome were drawn using R with gplots and ggplot2 packages. COGs categories of *T. kodakarensis* genes were extracted from 'The ArCOG database' (53) and used to draw the deviation graph with the plotrix R package.

RESULTS

Construction of gyrase-expressing *Thermococcus kodakarensis* strains

We transformed *T. kodakarensis* with the replicative plasmid pTNAg encoding *gyrA* and *gyrB* genes from *T. maritima*. The genes were expressed from *gyrAB* operon under the control of the constitutive archaeal promoter PhmtB. Single clone transformants that we named *T. kodakarensis* TKgyrAB strain (hereafter TKgyrAB) were readily obtained and the expression of the *gyrAB* was confirmed by RT-PCR (Supplementary figure S1). We also constructed a control strain carrying the empty vector (hereafter TKAg) and the strain expressing a mutated version of the gyrase where the catalytic tyrosine 119 was replaced by phenylalanine (hereafter TKY119F). This gyrase mutant was shown to bind DNA and ATP but is unable to generate double strand breaks necessary to introduce negative supercoiling in the DNA (54,55). We could not detect protein bands (90.5 kDa for GyrA and 72.5 for GyrB) corresponding to gyrase subunits on a SDS-PAGE gel indicating that in this experimental setup the gyrase was not expressed to very high levels compared to native proteins in TKgyrAB and TKY119F strains (Supplementary figure S2).

Plasmid DNA is negatively supercoiled in gyrase-expressing *T. kodakarensis*

Native plasmids of *Thermococcus* species are positively supercoiled *in vivo* presumably by the action of reverse gyrase (48). To investigate the impact of gyrase on DNA topology in *T. kodakarensis* we analysed the topology of plasmids isolated from the TKGyrAB, TKAg or TKY119F cultures using agarose gel electrophoresis. To facilitate these analyses, we introduced a smaller reporter plasmid pTPTK2 (5455 bp) into each of the three recombinant *Thermococcus* strains.

Plasmid topoisomers can be separated as single bands by one dimensional slow electrophoresis in agarose gels containing no DNA intercalating agent. In such conditions, the most highly supercoiled plasmids migrate the fastest forming a single front-band. We observed a broad range distribution of topoisomers in the two control strains as expected for plasmids with low level of positive supercoiling and only one fast-migrating major band of highly supercoiled DNA in the TKGyrAB strain (Supplementary figure S3A). This suggested that the gyrase was active in *T. kodakarensis*.

To confirm that the fast-migrating major band corresponded to negatively supercoiled DNA (in 1D electrophoresis, positively and negatively supercoiled topoisomers behave identically), we performed 2D agarose gel electrophoresis to determine the orientation (either positive or negative) of the supercoiling (56).

Initial 2D gel electrophoresis was performed in the absence of chloroquine in the first dimension to preserve the natural plasmid topology and with 7.5 μg/ml of chloroquine in the second dimension to determine the direction of supercoiling. This confirmed that the vast majority of the reporter plasmid was in an extreme negatively supercoiled state, in stark contrast to plasmid isolated from the control TKY119F strain (Figure 2A and C). The distribution of the topoisomers served as a reference to determine the number of supercoils introduced by chloroquine. Next, the plasmids were separated by adding chloroquine in both the first and the second dimension. This allowed us to determine the ΔLk for the major topoisomer in the gyrase expressing TKgyrAB strain (ΔLk = –8) as well as in the control strain TKY119F (ΔLk = –5) by taking into account the number of positive supercoils (+19) introduced by intercalation of chloroquine (Figure 2B and D). We used these values to calculate the mean supercoiling density from three independent experiments, after correction for the temperature effect (due to helix pitch increase) on the plasmid

topology (–10 supercoils) we obtained a superhelical density of +0.0077 ± 0.0016 for the control strain ($\sigma_{TKY119F}$) and –0.0327 ± 0.0016 for the strain containing the active DNA gyrase ($\sigma_{TKgyrAB}$).

Supercoiling density is used as a standardized measure of linking difference and can be compared between different organisms regardless of plasmid size or culturing conditions. The determined mean $\sigma_{TKY119F}$ was identical to that of the strain TKAg (Supplementary figure S3B) and it matched well with the reported native plasmid supercoiling level of *Thermococcus* sp. (48,57), thus suggesting that the inactive gyrase did not introduce topological changes in the reporter plasmid DNA (Figure 2E, supplementary table S3). The supercoiling density of gyrase-expressing *T. kodakarensis* matched that of Archaeoglobi archaea which are the only gyrase-encoding hyperthermophilic archaea (Figure 2E, asterisk, Supplementary table S3) (58). Remarkably, the increase in negative supercoiling density in TKgyrAB strain as compared to the control strain is more than 5-fold.

Collectively, the data show that active gyrase can be expressed in *Thermococcus kodakarensis* and that this organism can tolerate a substantial increase in negative supercoiling of its plasmid DNA. The data also suggest that the endogenous topoisomerases of *T. kodakarensis* with capacity to relax negatively supercoiled DNA (principally reverse gyrase and to a lesser extent Topo III and Topo VI) were outcompeted by gyrase.

The gyrase-expressing *T. kodakarensis* is sensitive to ciprofloxacin

We next asked if gyrase interacted with chromosomal DNA of *T. kodakarensis* and catalysed the double-stranded breaks required to introduce negative supercoiling. This can be tested indirectly by measuring the sensitivity of strains to the antibiotic ciprofloxacin. Ciprofloxacin binds to gyrase and DNA (59) and kills bacteria efficiently by stabilising the double-stranded break occurring during gyrase activity (60,61).

In contrast to what is reported for bacteria, we had to use relatively high concentrations of the drug to observe growth inhibition on plates (Figure 3). Since less than 10% of ciprofloxacin is degraded at the incubation temperature and time-scales we used (62), the low sensitivity of strain TKGyrAB to the drug may be explained by inefficient transfer of the drug across the membrane and/or very efficient efflux pumps. Intriguingly, we also observed a significant growth retardation of our control strains when exposed to ciprofloxacin indicating that the drug interferes with essential process(es) in *T. kodakarensis*. The figure 3 shows the result of the antibiogram test on phytagel plates. The cells were plated on non-selective medium to make sure that we detected the toxic effect of the ciprofloxacin (gyrase-DNA covalent adducts) against the chromosome and not against the plasmid that carried the selective marker. The assay repeatedly ($n = 4$) showed higher sensitivity of TKGyrAB strain to the ciprofloxacin compared to controls, consistent with the formation of toxic gyrase-DNA covalent adducts on the chromosome.

Global transcriptional response to gyrase expression in *Thermococcus kodakarensis*

To understand better how *T. kodakarensis* cells cope with the presence of artificial negative supercoiling activity we investigated genome-wide transcriptional responses in the three recombinant strains.

We performed RNA-seq experiments on biological replicates ($n = 4$) of exponentially growing cells and quantified differential transcript abundance in TKgyrAB versus TKAg, TKY119F versus TKAg and TKgyrAB versus TKY119F (Supplementary figure S4). We reasoned that the latter analysis would, in principle, allow us to identify the subset of genes enriched for those responding specifically to the negative supercoiling activity of the gyrase while TKY119F versus TKAg comparison would yield genes responding to the burden of gyrase heterologous expression and its DNA binding activity. In agreement with this hypothesis, only few differentially expressed genes (DEGs) are shared between TKgyrAB vs TKY119F (genes responding to negative supercoiling) and TKY119F versus TKAg (genes responding to gyrase expression burden and DNA binding) or in all three comparisons (Supplementary figure S5).

We first compared the transcriptional profile of the genes encoded on the empty vector and the two plasmids carrying gyrase genes. This was particularly interesting because, in contrast to the chromosome where the local DNA topology status was not known, we knew that the plasmid encoding active gyrase was negatively supercoiled and thus we could measure how negative supercoiling affected gene transcription. Negative supercoiling facilitates DNA melting and thereby promotes transcription (63,64). We therefore anticipated an increase in expression of plasmid-borne genes when comparing TKgyrAB vs TKY119F transcriptional profiles. Intriguingly, however, none of the genes (including the *gyrAB* operon) were differentially expressed (Supplementary table S4). Interestingly, the expression of inactive gyrase was not neutral, three out of five plasmid-encoded genes were significantly downregulated suggesting that gyrase might bind within or in the vicinity of these genes and interfere with their transcription. It has to be noted though that the interpretation of these data needs to be taken with precaution since the precise plasmid copy number in each strain is not known.

We next quantified differential transcript abundance for chromosomal genes. The expression of inactive gyrase alone modified significantly (Padj < 0.05) the expression of 143 genes (fold change > 1.25) most of which (80%) were downregulated in agreement with the hypothesis that the DNA binding activity of gyrase alone impedes transcription (Figure 4 A). This suggested a dominant negative effect that could be explained by the formation of a stable complex in which DNA is wrapped around the CDT domain of the gyrase heterotetramer. Without the possibility to cut the DNA, it is possible that the catalytically dead gyrase would stay bound to DNA thus forming a mechanical barrier for the passage of the RNA polymerase. When active gyrase is expressed, 410 genes were affected which corresponds to ∼18% of total number of annotated genes. Finally, by comparing TKgyrAB and TKY119F expression profiles we

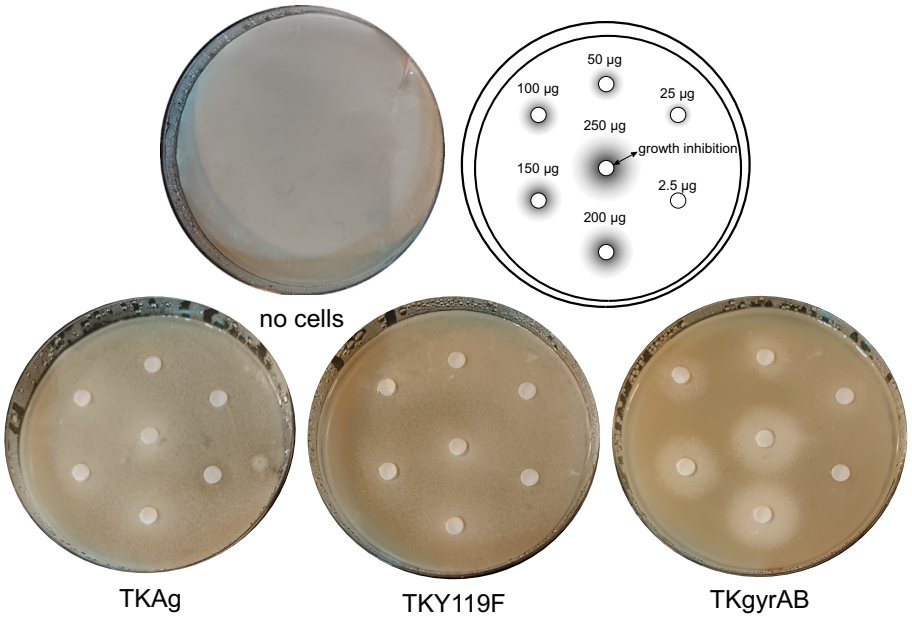


Figure 3. Gyrase expression induces ciprofloxacin sensitivity in *T. kodakarensis Thermococcus* cells from overnight cultures were mixed with 0.18% (w/v) phytagel and colloidal sulfur containing culture medium and were spread as overlay onto 1% phytagel plates. The plate cartoon indicates the position of the paper discs and the quantity of ciprofloxacin used. The cell-containing overlay becomes transparent when incubated at 85°C due to consumption of colloidal sulphur during *T. kodakarensis* growth. Growth inhibition can therefore be detected as a white area on the plate.

identified 205 DEGs which specifically responded to negative supercoiling activity of the gyrase. The comparison of the relative transcript abundances for these genes using *Z*-score scaling shows an opposite tendency between the gyrase expressing strain and the two control strains (Supplementary figure S6). We have therefore conducted further analyses on this set of genes which we named SRGs for supercoiling-responding genes.

SRGs were distributed throughout the entire chromosome of *T. kodakarensis* without obvious bias with respect to GC skew or transcription direction (Figure 4, B). We did not detect SRGs-specific GC content in the promoter region of SRGs or bias in the size of the corresponding transcripts (Supplementary figure S9). We noticed however, hotspots of deregulated genes in vicinity of ribosomal RNA genes. These regions would be expected to be targeted by the gyrase because the flanking regions of heavily transcribed ribosomal RNA genes are rich in supercoiled DNA (65).

The most highly deregulated SRGs encode the archaellum components (FC $\sim$ 54) and the genes involved in chemotaxis which are both highly upregulated (FC $\sim$ 6). An opposite downregulation trend was observed for genes assigned to energy production and conversion, nucleotide transport and metabolism and cell cycle control, cell division and chromosome partitioning functional categories (Supplementary figure S7). Among 11 DNA-repair related genes (66) only TK0784 encoding a homologue of XPD was significantly deregulated ($\log_2$FC = 0,33) suggesting that gyrase expression does not induce high levels of DNA damage. Intriguingly, however, the genes encoding Topo VI or Topo III were not deregulated although these enzymes are known to relax negatively supercoiled DNA *in vitro* (Supplementary Table S5). A notable exception was the reverse gyrase-encoding gene showing significant upregulation ($P_{adj} = 0.007$) albeit with a low fold change (FC = 1.2). We next looked into the transcript level of gyrase and reverse gyrase which, although an imperfect proxy of protein quantity, can give some insight into relative protein abundances in cells. To be able to compare the read counts within each sample we converted them into fragments per kilobase million (FPKM, Supplementary table S6). Assuming that all the transcripts are translated into functional protein, this shows that reverse gyrase, Topo VI and Mini-A (distant homolog of type IIB topoisomerases, (67)) are present in comparable levels in *T. kodakaransis*. In GyrAB strain average FPKM values are 1.5 for GyrA, 0.6 for GyrB and 0.47 for reverse gyrase suggesting that the gyrase is the most abundant topoisomerase in the GyrAB strain. From these data, we can speculate that the observed upregulation of reverse gyrase in the GyrAB strain was insufficient to counteract the gyrase negative supercoiling activity as suggested by the topology of the plasmids.

The above results demonstrated that gyrase introduction in *T. kodakarensis* provoked a genome-wide but, in most cases, mild deregulation of gene expression. The data also suggested that the gyrase-derived negative supercoiling was not handled by the endogenous topoisomerases.

To understand better the molecular bases underpinning the observed transcriptional response, we compared our data with previously published analyses of *T. kodakarensis* transcriptomes. Among six available datasets one caught our attention as, similar to our study, the *fla* (archaellum) and *che* (chemotaxis) operons were the most highly deregulated (68). Sanders and colleagues studied the transcriptional profile of *T. kodakarensis* that either could not build multimeric chromatin particles (strain TS620,

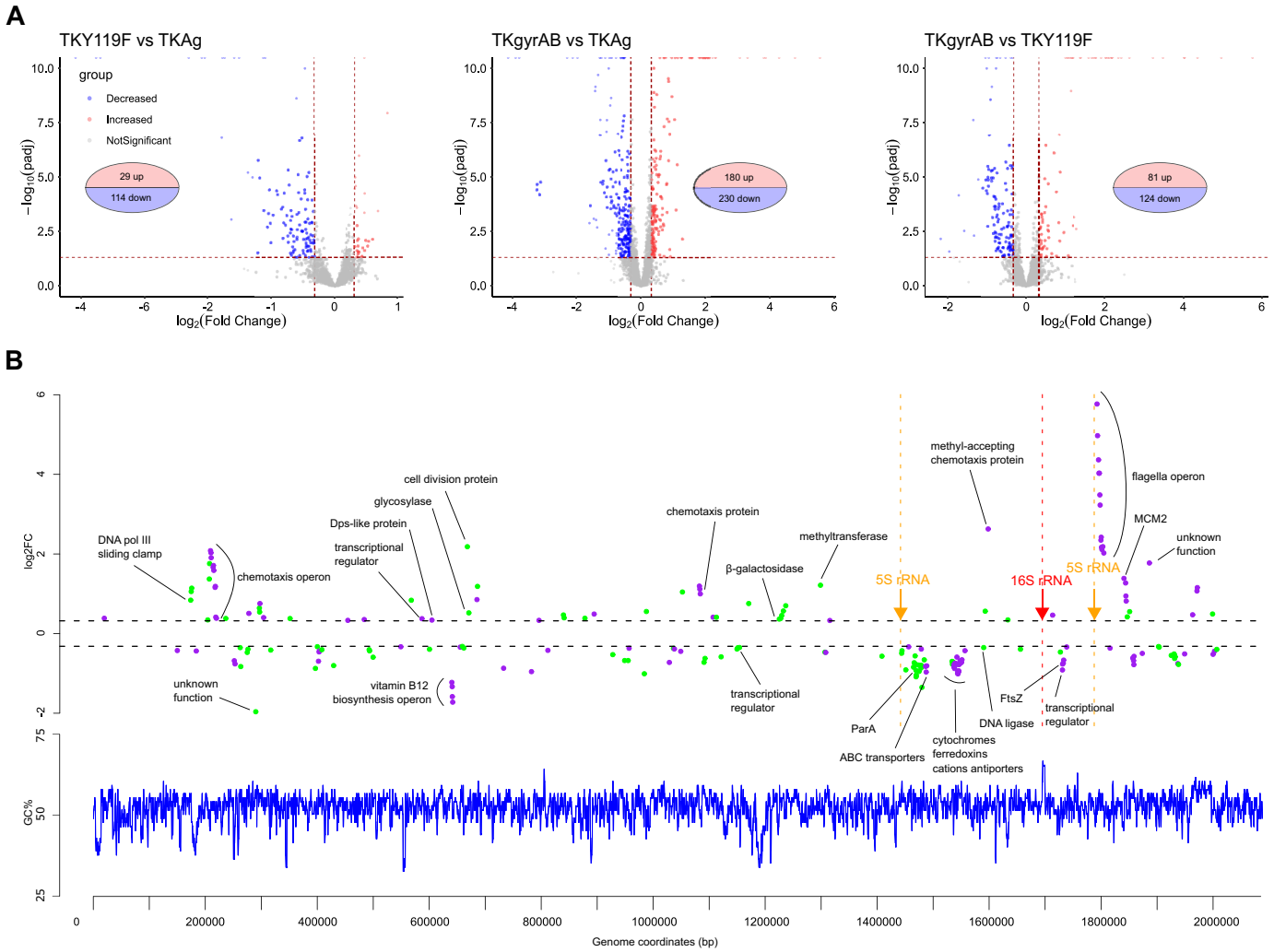


Figure 4. The impact of gyrase on transcription in *T. kodakarensis* (**A**) Volcano plots showing significantly deregulated genes. The two vertical and the horizontal dashed lines indicate the threshold values of ±1.25 fold change and Padj > 0.05, respectively. The blue and red dots correspond to downregulated and upregulated genes, respectively. Each plot corresponds to a pairwise comparison of transcriptomes from strains indicated on the top. (**B**) Distribution of the SRGs on the chromosome of *T. kodakarensis* TS599. Top panel shows the distribution of DEGs on the chromosome of *T. kodakarensis*. Points correspond to individual DEGs and their colour indicates the orientation of transcription whereby green indicates antisense expressed genes and violet indicates sense expressed genes. The position of ribosomal genes is indicated by arrows. The annotated function of SRGs involved in DNA transactions and cell division (COG categories L, K, D and B) as well as those of outliers is indicated. The lower panel shows the distribution of the GC content along the chromosome.

ΔHTkB HTkA^{G17D}) or relied only on Histone B for building histone polymers (strain TS622, HTkB^{WT} HTkA^{G17D}). Interestingly, out of top 30 upregulated SRGs, at least 26 were downregulated in strains TS620 and TS622 (Figure 5A). Most of these genes belonged to *fla* and *che* operons but we also identified two non-operonic chemotaxis genes (TK0156 and TK2147), one gene annotated as AAA + ATPase (TK1139) and a predicted FprA family A-type flavoprotein electron transfer protein (TK1605). When we extended this analysis to all SRGs (298 genes with $P_{adj} < 0.05$) we did not find any correlation with transcriptomes of TS620 and TS622 (Supplementary figure S8). Together, these data suggested that only the most highly upregulated SRGs are sensitive to both gyrase induced negative supercoiling and chromatin structural defect. To further examine the molecular basis for such behaviour we mapped transcriptional start sites (TSS) for anti-correlated SRGs based on experimental data from Jäger and coll. (69). The six identified intergenic sequences were subjected to MEME analysis (70) which revealed the presence of a 23 bp common motif (consensus sequence TTTGTGTABSTGBTTATGTAGGT) present in one copy or in two copies for *fla* operon. The motif, mostly located ∼35 bp upstream of TSS, does not resemble consensus promoter motifs of *T. kodakarensis* which typically have a B recognition element (BRE) followed by a TATA-box ∼33 and ∼23 bp, respectively, upstream from transcription initiation (69) (Figure 5B). A search for the motif in whole *T. kodakarensis* genome using FIMO (70) retrieved the six already identified motifs but no additional high scoring hits.

This analysis thus uncovered a common sequence motif found exclusively in the promoter region of the most highly

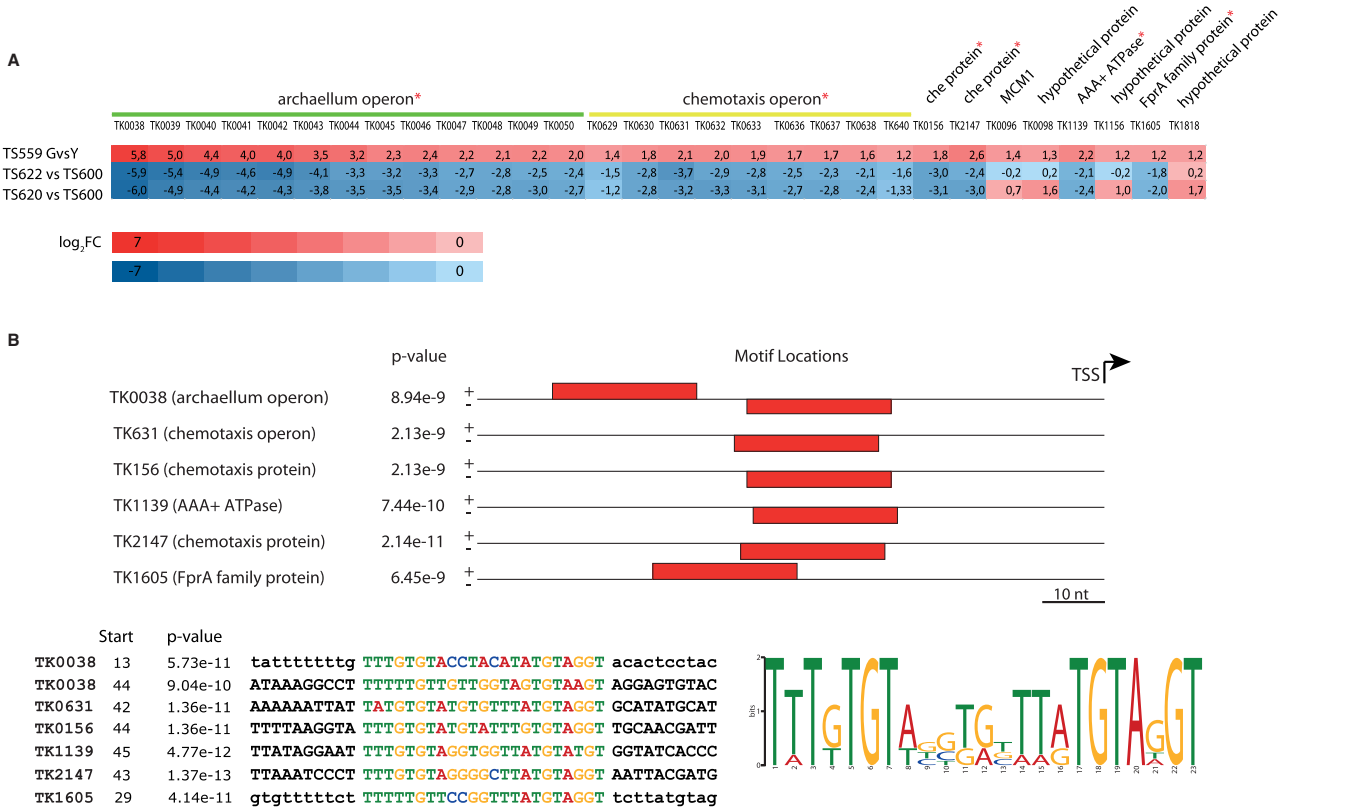


Figure 5. Top upregulated SRGs carry specific sequence motif in their promoter region. (**A**) Differentially expressed genes reacting to gyrase supercoiling activity or chromatin defect. Rectangles correspond to the top 30 upregulated SRGs and their colour corresponds to the $\log_2$FC values as indicated by the scale. The $\log_2$FC values of the genes/operons indicated by an asterisk are systematically anti-correlated between the gyrase expressing strain and the *T. kodakarensis* TS620 and TS622 strains. (**B**) Upper panel shows that the anti-correlated genes/operons invariably contain a conserved 23 bp sequence motif, indicated as red rectangle, upstream of the transcription start site (TSS). The lower panel shows the alignment of the 23 bp sequences on the left and the corresponding sequence logo on the right.

upregulated SRGs, and further suggests that the dysregulation of these genes might be the consequence of alterations of chromatin structure induced by the negative-supercoiling activity of the gyrase.

Impact of gyrase on *Thermococcus kodakarensis* growth

To assess the impact of gyrase on *T. kodakarensis* growth we first monitored the growth kinetics of the three strains in batch cultures and at optimal growth conditions by direct cell counting. The recorded growth curves exhibited a typical sigmoidal shape and were overall similar (Figure 6A). The specific growth rate at the exponential phase was lower for TKgyrAB and TKY119F as compared to TKAg control strain, however, the slopes were not significantly different ($P = 0.34$) indicating that the observed differences are not significant.

Next, the cell shape and size as well as DNA content were assessed using light microscopy coupled with image analysis. This revealed that the exponential phase TKgyrAB cells had the typical irregular round shape as the control cells and contained DNA. However, the measurement of cell area revealed that the gyrase expressing strains were, on average, smaller than the TKAg strain and this difference was statistically significant ($P = 8.1 \times 10^{-70}$, Figure 6B, Supplementary figure S10). The phenotype was the most pronounced for TKGyrAB cells which (assuming that the cells are perfectly spherical) exhibited in average 35% less volume compared to TkAg control. Despite being smaller, the TKGyrAB cells did not exhibit a filamentous phenotype indicating they divided at a normal rate (Supplementary Figure S10).

We next tested whether the gyrase expression becomes toxic during longer culturing. If so, we expected that the gyrase expression plasmid would be gradually lost in non-selective culture conditions. We therefore quantified the gyrase expressing plasmid in the population by counting colony forming units (CFU) under selective versus non-selective conditions over 84 generations (14 subcultures). Plasmid was stably maintained in the TKgyrAB strain while about 50% plasmid loss was observed for strain TKY119F at the final stage of the experiment (Figure 6C). To ensure that the experiment was performed with cells that contained active gyrase we monitored pTPTK2 topology using 1D agarose gel electrophoresis (Supplementary figure S11). The amount of the extreme supercoiled form of pTPTK2 remained constant throughout the experiment indicating no loss of gyrase activity during prolonged culturing of *Thermococcus*.

Collectively, the data show that the gyrase is remarkably well tolerated by *T. kodakarensis*.

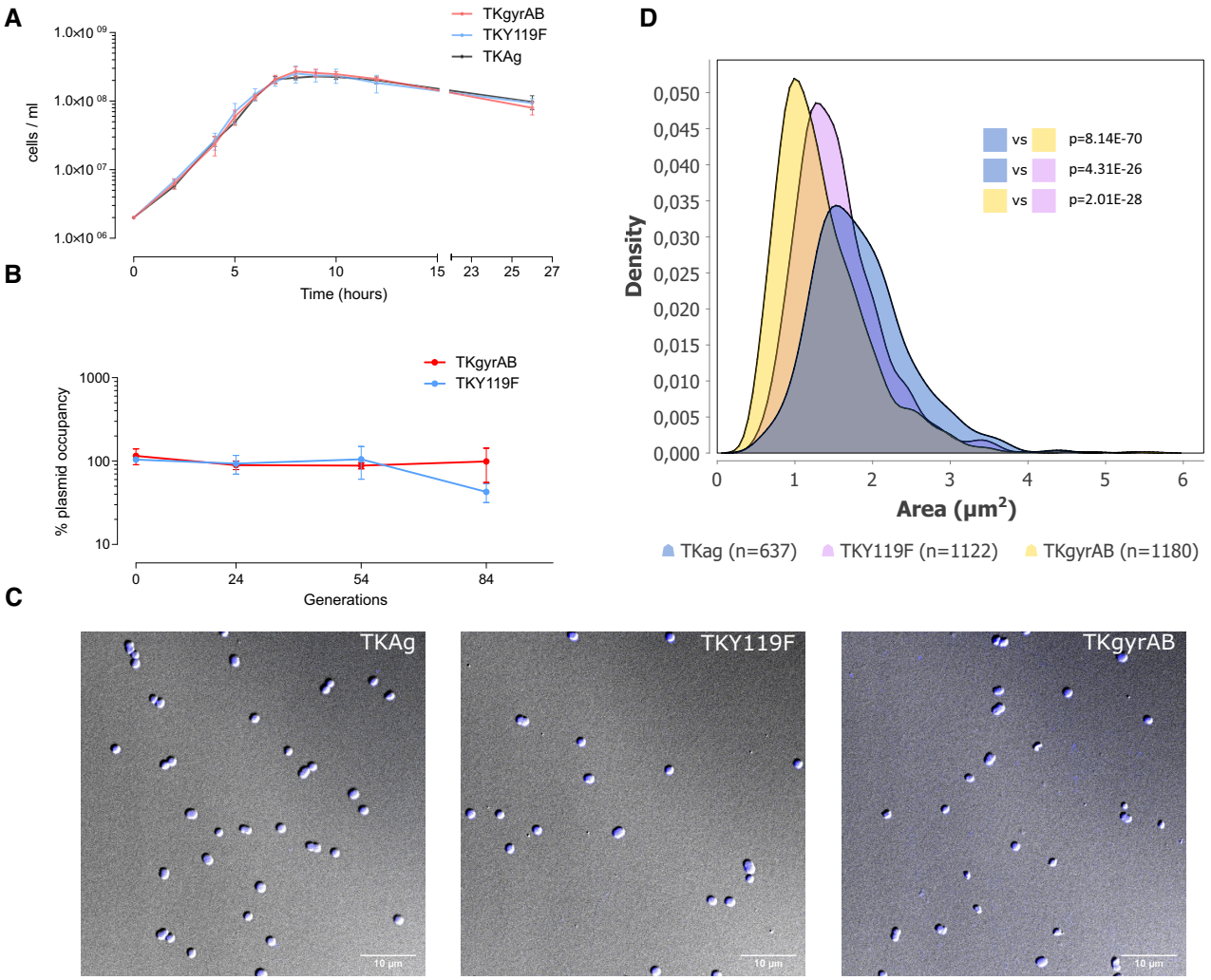


Figure 6. Expression of active DNA gyrase is not toxic for *T. kodakarensis*. (**A**) Growth of the three strains was monitored by direct cell counting using a Thoma cell counting chamber. The specific growth rates (number of divisions h^{-1}) of the control (TKY119, TKAg) and gyrase expressing strain TKgyrAB) were 2.11 ± 0.27 h^{-1} TKAg, 1.45 ± 0.35 h^{-1} TKY119 and 1.81 ± 0.37 h^{-1} TKgyrAB. Specific growth rates were calculated from the slope of the linear portions of the curves according to the equation $\mu = dY/dt$, where t is time and Y is the cell density. Mean values of Y were used for the calculation. Error bars represent the standard deviation, $n = 3$. The differences between the slopes are not significant, $P = 0.34$, ANCOVA two-tailed test. (**B**) Measurement of plasmid loss over 84 generations corresponding to 14 subcultures. Each point corresponds to the ratio of CFUs grown on plates under non-selective or selective medium. The experiment was done in triplicate and the bars correspond to standard deviation from the mean value. (**C**) Representative micrographs of *T. kodakarensis* cells harvested at the exponential growth stage. The cells were stained with Hoechst dye and DIC and fluorescent images were superposed. (**D**) Density plot of cell area determined from DIC microscopy images. The number of analysed cells is indicated below the graph.

DISCUSSION

In this study, we demonstrate that it is possible to introduce active bacterial gyrase in the hyperthermophilic archaeon *T. kodakarensis*, a cellular system that naturally lacks negative supercoiling activity. We further show that, as in bacteria, the gyrase became the dominant topoisomerase converting positively supercoiled plasmids into highly negatively supercoiled DNA. Transcriptomic analyses revealed mild deregulation of hundreds of genes including induction of stress – related flagellar (archaellum) and chemotaxis systems. The analysis of the top 30 upregulated genes (including *fla* and *che* operons) revealed the presence of a conserved 23 bp sequence motif in their promoter region. These genes were also systematically downregulated in *T. kodakarensis* strains carrying mutations in genes encoding histones. Reverse gyrase was the only topoisomerase of *T. kodakarensis* for which the expression was altered (slightly upregulated) in response to the negative supercoiling activity of gyrase. Despite global-scale alterations of its cellular context, the *T. kodakarensis* growth rate was not affected, suggesting that critical DNA-templated processes were not compromised by gyrase activity.

This was an unexpected result because the negative supercoiling activity of the gyrase should, in principle, interfere with essential processes such as transcription or DNA replication. In particular, gyrase relaxes positive supercoils accumulating ahead of transcribing RNA polymerase (18), a task predicted to be accomplished by Topo VI in archaea (71). Moreover, negative supercoiling facilitates DNA melting which in turn facilitates promoter firing (64) but, at high

temperatures, also exposes ssDNA to heat-induced damage (72,73). In spite of these potential threats to genome stability and expression, *T. kodakarensis* expressing gyrase tolerated an approximately five-fold increase in the negative supercoiling density of plasmid DNA as compared to the natural plasmid DNA topology. Granted, plasmid DNA topology does not necessarily recapitulate chromosomal DNA topology, this observation is still quite impressive when put in perspective with the life-limiting tolerance level of bacteria. *Streptococcus pneumoniae* can only tolerate increases in negative supercoiling levels of up to 1.4-fold (≈41% increase for plasmid DNA) (74) and *E. coli* of up to 1.1 fold (≈14% increase for plasmid and chromosomal DNA) (75). Besides, the RNA-seq data do not show notable induction of genes involved in DNA repair thus suggesting that the gyrase expression does not result in massive DNA damage in *T. kodakarensis*.

Why is gyrase so well tolerated by an organism in which the cellular DNA and machineries did not co-evolve to accommodate pervasive negative supercoiling activity? We suggest that during evolution, and in *T. kodakarensis,* the wrapping of genomic DNA in nucleosome-like structures may have facilitated the establishment of the gyrase in archaeal cells. In Euryarchaea, histones have evolved as most abundant chromatin proteins (76,77) and *T. kodakarensis* has particularly dense histone coverage approaching 100% (69,78–80). Commonly, archaeal histone dimers assemble into tetramers as minimal nucleosomal units that wrap ∼ 60 bp of DNA (81) but in some archaea, including *T. kodakarensis*, the tetramers can be extended via incorporation of additional dimers into particles of variable size that wrap up to 480 bp of DNA in negatively constrained supercoils (78,82,83). This flexible chromatin structure may participate in adaptive responses of *T. kodakarensis* by restructuring the existing pool of histones such that the local gyrase-induced increase of negative supercoiling would be efficiently absorbed. In line with this idea, we discovered that the most highly upregulated SRGs are almost systematically downregulated in *T. kodakarensis* strains carrying mutations in histone genes. Remarkably, the anti-correlated transcriptional response is always associated with a 23 bp motif occurring about 35 bp upstream of transcription start site of these SRGs. The function of this motif is currently unclear but we can speculate that its presence may render the promoter region particularly sensitive to alterations of DNA topology. Along the same line, it is interesting to note that artificial chromatinization of the *E. coli* genome with archaeal histones resulted in downregulation of gyrase genes as a part of the adaptive cellular response (84). Together, these observations gathered in artificial settings suggest that both organisms tolerate well the introduction of major modellers of DNA topology and react by balancing chromatin structure and DNA supercoiling to achieve a DNA geometry necessary to sustain life.

Once established, the gyrase became fixed in many archaeal lineages to the point where it has become an essential protein in present-day archaea. Indeed, early studies showed that gyrase-targeting drugs such as novobiocin and ciprofloxacin inhibited growth of different archaea including methanogens, halophiles and thermoacidophiles (85). These experiments also established that gyrase was responsible for introducing most, if not all of the negative supercoils in plasmid DNA molecules (85,86). Notably, novobiocin treatment of *Halobacterium halobium* cultures stopped DNA replication specifically and instantaneously indicating, albeit indirectly, that gyrase acts upon chromosomal DNA during the elongation step of DNA replication (85). Collectively, these findings suggest that the negative supercoiling activity of the gyrase was positively selected in the course of evolution, but the selective advantage conferred by this feature remains to be established and future studies of *in vivo* functions of archaeal gyrases should bring some insight. In Bacteria, gyrase-controlled negative supercoiling is instrumental for quick adaptation of the cellular protein repertoire to changing environmental conditions (20) but whether such mechanisms operate in gyrase-encoding Archaea is not known. Indirect evidence points to involvement of negative DNA supercoiling in gene expression control in extreme halophiles. In these organisms, a plasmid-encoded *gyrB* gene and chromosomally encoded *bop* gene (encoding bacteriorodopsin) were strongly induced (up to 20-fold) by DNA relaxation in novobiocin-treated cultures, a drug that inhibits gyrase activity (87,88). More recently, a global transcriptome analysis was reported for novobiocin-treated *Halobacterium* species (89). The expression of many genes was affected including the upregulation of gyrase, topoisomerase VI and topoisomerase I expression indicating the involvement of these enzymes in regulation of the DNA supercoiling levels in this organism (89). However, to what extent the gyrase controls the chromosomal supercoiling and how its activity is coordinated with other archaeal topoisomerases, histones and NAPs is currently unknown.

Related to that, it is noteworthy that reverse gyrase was the only topoisomerase of *T. kodakarensis* that responded to the negative supercoiling activity of the gyrase. Reverse gyrase is a Topo I enzyme and the only topoisomerase capable of supercoiling DNA positively (32). Through this activity, reverse gyrase can remove negative supercoiling (32,33,90,91). Remarkably, reverse gyrase is found specifically in thermophilic organisms and its deletion in *T. kodakarensis* and *Pyrococcus furiosus* is lethal at 93 and 95°C, respectively (34,35). It was initially suggested that reverse gyrase prevents thermal denaturation of the double helix by introducing positive supercoils in chromosomal DNA (32–34). However, the idea that positively supercoiled DNA, in spite of its stabilising effect, is not essential for a hyperthermophilic lifestyle was put forward some time ago based on the observation that thermophilic *Thermotoga* bacteria and gyrase-encoding hyperthermophilic Archaeoglobi archaea contain negatively supercoiled plasmids (40,58). Later studies reported the involvement of this topoisomerase in DNA repair (92–95). Recently, however, a study of gyrase-less *Saccharolobus* (formerly *Sulfolobus*) *solfataricus* reported that reverse gyrase is involved in homeostatic control of DNA supercoiling mainly based on the finding that the protein abundance increased about two-fold when cells were exposed to supraoptimal temperatures and the enzyme was more active *in vitro* (91). Our data show that 1.2-fold (at the mRNA level) upregulation of reverse gyrase in *T. kodakarensis* is insufficient to restore natural DNA topology and suggest that, at least in *Thermococcus* and in these artifi-

cial conditions, this topoisomerase is not essential for homeostatic control of DNA topology.

In contrast, we were intrigued to find that the transcription profile of plasmid-encoded genes in the gyrase-expressing *T. kodakarensis* seems not to be modified. The most clearly identified effect of supercoiling on transcription initiation results from the requirement of RNA polymerase to open the double helix in order to gain access to the template strand. Negative supercoiling facilitates melting of the double helix and a strong regulatory effect of negative supercoiling upon gene transcription is well recorded in bacteria (21,65). In hyperthermophiles, which have relaxed or slightly positively supercoiled DNA, it is thought that the elevated growth temperature may replace negative supercoiling as source of melting energy (96,97) and we therefore expected that the gyrase–induced negative supercoiling and high temperature would synergistically activate the expression of plasmid-borne genes in *T. kodakarensis*. The fact that this is not so suggests that *T. kodakarensis* is naturally equipped to allow transcription from topologically different DNA templates. This is reminiscent of observations made by Bell and colleagues who studied the effect of temperature and template topology on expression from an archaeal ribosomal RNA promoter using a highly purified *in vitro* system from crenarchaeon *Sulfolobus* (63). They found that, in marked contrast to characterised bacterial and eukaryal systems, DNA template topology had negligible effect on transcription levels at 78°C, the optimal growth temperature for *Sulfolobus*. In another study, Hethke *et al.* used the cell-free transcription system of *Pyrococcus furiosus,* a euryarchaeon closely related to *Thermococcus*, to study the effect of DNA template topology on expression from the *gdh* promoter at 70 and 90°C (73). They found that at both temperatures negatively supercoiled DNA was the preferred template compared with relaxed DNA and that positive supercoiling deteriorates the template activity of DNA. It is unclear whether the differences between *P. furiosus* and *Sulfolobus* transcription systems come from the type of promoter used or from the intrinsic properties of their transcription machineries (73). We can now use the gyrase-expressing *T. kodakarensis* to perform similar studies *in vivo* and thus study the relationship between supercoiling and gene expression in more natural settings.

Would global negative supercoiling be advantageous to a hyperthermophile at lower, suboptimal temperatures? The experiments described by Bell and colleagues suggest so since the *Sulfolobus* system was unable to transcribe relaxed or positively supercoiled templates at 48°C. They further highlighted that, in response to cold shock, hyperthermophiles rapidly reduce their plasmid linking number (48,98) raising the possibility that global regulation of DNA superhelical state *in vivo* represents an effective mechanism for ensuring continued gene expression after drastic changes in temperature of the environment. Building upon these findings, more than 20 years ago, P. López-García proposed that archaea may have improved their adaptability to mesophily by importing the gyrase from bacteria (99). This idea was reinforced later on by the finding that the vast majority of gyrase-encoding monophyletic group II Euryarchaea are mesophiles in spite of their thermophilic origin (13,26,28). These archaea possess histones leading to the proposal that the acquisition of the gyrase may have had a synergistic effect on DNA-dependent processes in these organisms, with associated changes in transcriptional patterns thus contributing to bacterial-like progressive adaptation to lower temperatures (28). The gyrase-expressing *T. kodakarensis* now offers the possibility to test this evolutionary hypothesis and ultimately understand why several archaeal lineages became addicted to gyrase.

DATA AVAILABILITY

The RNA-seq data reported in this article are available in ArrayExpress (https://www.ebi.ac.uk/arrayexpress/) database and can be accessed with E-MTAB-10799 accession number. Processed tables, which were modified to homogenize gene names, are available upon request.

SUPPLEMENTARY DATA

Supplementary Data are available at NAR Online.

ACKNOWLEDGEMENTS

We would like to thank fellow scientists who helped us along the way. Olivier Namy, Isabelle Hatin and Pauline François for help with design of RNA-seq experiments and data handling. Laura Baranello and Jan Grosser, Charles Dorman and Virginia Lioy for their feedback on our data and constructive discussions. Clémence Lauden for help with plotting under R studio. We thank Aurore Gorlas and Evelyne Marguet for advice regarding 2D electrophoresis and Augustin Degaugue for technical support with plasmid loss experiments.

We acknowledge the High-throughput sequencing facility of I2BC (Centre de Recherche de Gif – http://www.i2bc.paris-saclay.fr/) for its sequencing and bioinformatics expertise with particular thanks to Yan Jaszczyszyn, Erwin van Dijk and Céline Hernandez for their technical advice and help with data analysis.

The present work has benefited from Imagerie-Gif core facility, supported by l'Agence Nationale de la Recherche (ANR-11-EQPX-0029/Morphoscope, ANR-10-INBS-04/FranceBioImaging; ANR-11-IDEX-0003-02/SaclayPlant Sciences. Many thanks to Romain Le Bars for helping us with microscopy data acquisition and analysis.

Author contributions: P.V. performed the experiments and analysed data. VdC, E.V. and J.O. analysed data. T.B, R.C. and P.F. designed the research. J.O. funded the research. P.V. and T.B. wrote the paper. All the authors read and approved the manuscript.

FUNDING

P.V. received the 3 year PhD scholarship from the French Ministry of Higher Education and Research; R.C. and VdC postdoctoral fellowship was funded by the European Research Council under the European Union's Seventh Framework Program [FP/2007–2013]/Project EVO-MOBIL - ERC Grant Agreement [340440]. Funding for open access charge: CNRS.

Conflict of interest statement. None declared.

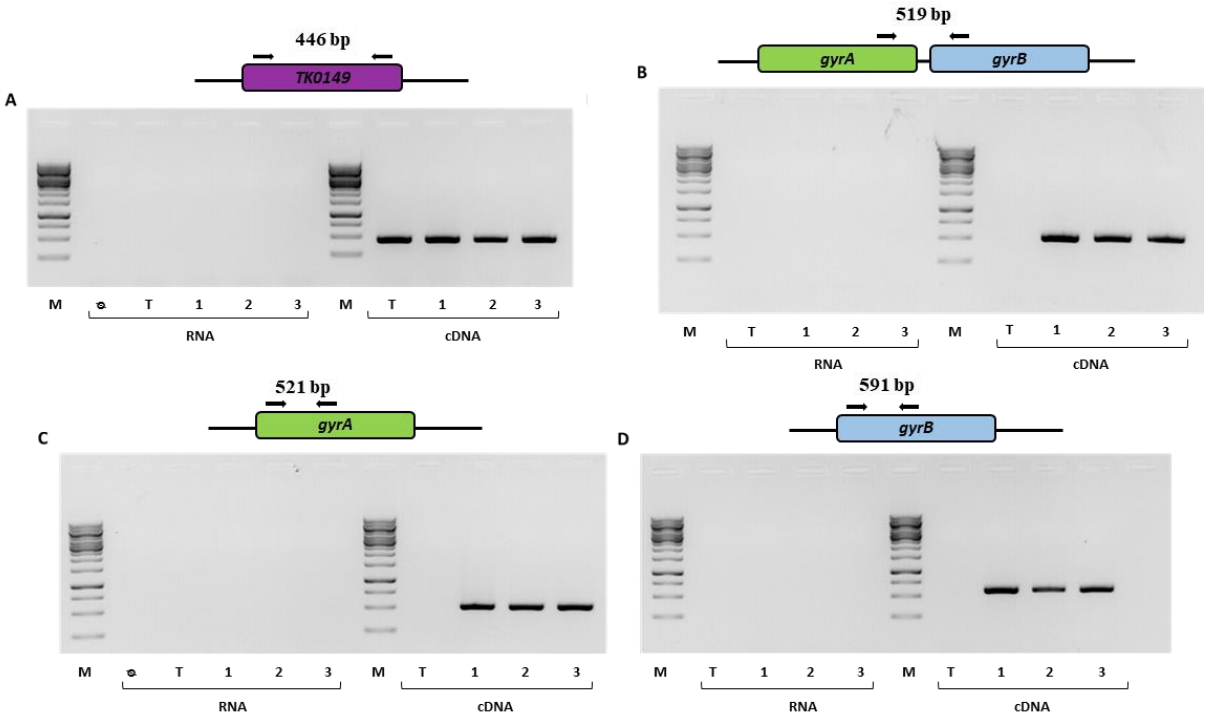
REFERENCES

1. Watson,J.D. and Crick,F.H.C. (1953) Molecular structure of nucleic acids: a structure for deoxyribose nucleic acid. *Nature*, **171**, 737–738.
2. Liu,L.F. and Wang,J.C. (1987) Supercoiling of the DNA template during transcription. *Proc. Natl. Acad. Sci. U.S.A.*, **84**, 7024–7027.
3. Postow,L., Crisona,N.J., Peter,B.J., Hardy,C.D. and Cozzarelli,N.R. (2001) Topological challenges to DNA replication: Conformations at the fork. *Proc. Natl Acad. Sci. U.S.A.*, **98**, 8219–8226.
4. Gilbert,N. and Allan,J. (2014) Supercoiling in DNA and chromatin. *Curr. Opin. Genet. Dev.*, **25**, 15–21.
5. Wang,J.C. (2002) Cellular roles of DNA topoisomerases: a molecular perspective. *Nat. Rev. Mol. Cell Biol.*, **3**, 430–440.
6. Schoeffler,A.J. and Berger,J.M. (2008) DNA topoisomerases: harnessing and constraining energy to govern chromosome topology. *Quart. Rev. Biophys.*, **41**, 41–101.
7. Forterre,P. (2011) Introduction and Historical perspective. In: *DNA Topoisomerases and Cancer*. Humana press, Springer-Verlag, pp. 1–52.
8. Vos,S.M., Tretter,E.M., Schmidt,B.H. and Berger,J.M. (2011) All tangled up: how cells direct, manage and exploit topoisomerase function. *Nat. Rev. Mol. Cell Biol.*, **12**, 827–841.
9. Bush,N.G., Evans-Roberts,K. and Maxwell,A. (2015) DNA topoisomerases. *EcoSal Plus*, **6**, https://doi.org/10.1128/ecosalplus.ESP-0010-2014.
10. Seol,Y. and Neuman,K.C. (2016) The dynamic interplay between DNA topoisomerases and DNA Topology. *Biophys. Rev.*, **8**, 221–231.
11. McKie,S.J., Neuman,K.C. and Maxwell,A. (2021) DNA topoisomerases: Advances in understanding of cellular roles and multi-protein complexes via structure-function analysis. *Bioessays*, **43**, 2000286.
12. Forterre,P., Gribaldo,S., Gadelle,D. and Serre,M.-C. (2007) Origin and evolution of DNA topoisomerases. *Biochimie*, **89**, 427–446.
13. Forterre,P. and Gadelle,D. (2009) Phylogenomics of DNA topoisomerases: their origin and putative roles in the emergence of modern organisms. *Nucleic Acids Res.*, **37**, 679–692.
14. Gellert,M., Mizuuchi,K., O'Dea,M.H. and Nash,H.A. (1976) DNA gyrase: an enzyme that introduces superhelical turns into DNA. *Proc. Natl. Acad. Sci. U.S.A.*, **73**, 3872–3876.
15. Drlica,K. (1992) Control of bacterial DNA supercoiling. *Mol. Microbiol.*, **6**, 425–433.
16. Zechiedrich,E.L., Khodursky,A.B., Bachellier,S., Schneider,R., Chen,D., Lilley,D.M. and Cozzarelli,N.R. (2000) Roles of topoisomerases in maintaining steady-state DNA supercoiling in Escherichia coli. *J. Biol. Chem.*, **275**, 8103–8113.
17. Lal,A., Dhar,A., Trostel,A., Kouzine,F., Seshasayee,A.S.N. and Adhya,S. (2016) Genome scale patterns of supercoiling in a bacterial chromosome. *Nat. Commun.*, **7**, 11055.
18. Sutormin,D., Rubanova,N., Logacheva,M., Ghilarov,D. and Severinov,K. (2019) Single-nucleotide-resolution mapping of DNA gyrase cleavage sites across the Escherichia coli genome. *Nucleic Acids Res.*, **47**, 1373–1388.
19. Hardy,C.D., Crisona,N.J., Stone,M.D. and Cozzarelli,N.R. (2004) Disentangling DNA during replication: a tale of two strands. *Philos. Trans. R. Soc. Lond., B, Biol. Sci.*, **359**, 39–47.
20. Dorman,C.J. and Dorman,M.J. (2016) DNA supercoiling is a fundamental regulatory principle in the control of bacterial gene expression. *Biophys. Rev.*, **8**, 209–220.
21. Martis,B, S., Forquet,R., Reverchon,S., Nasser,W. and Meyer,S. (2019) DNA supercoiling: an ancestral regulator of gene expression in pathogenic bacteria? *Comput Struct Biotechnol J*, **17**, 1047–1055.
22. Westerhoff,H.V., O'Dea,M.H., Maxwell,A. and Gellert,M. (1988) DNA supercoiling by DNA gyrase. *Cell Biophys.*, **12**, 157–181.
23. Hsieh,L.S., Rouviere-Yaniv,J. and Drlica,K. (1991) Bacterial DNA supercoiling and [ATP]/[ADP] ratio: changes associated with salt shock. *J. Bacteriol.*, **173**, 3914–3917.
24. Adam,P.S., Borrel,G., Brochier-Armanet,C. and Gribaldo,S. (2017) The growing tree of Archaea: new perspectives on their diversity, evolution and ecology. *ISME J*, **11**, 2407–2425.
25. Garnier,F., Couturier,M., Débat,H. and Nadal,M. (2021) Archaea: a gold mine for topoisomerase diversity. *Front. Microbiol.*, **12**, 661411.
26. Raymann,K., Forterre,P., Brochier-Armanet,C. and Gribaldo,S. (2014) Global phylogenomic analysis disentangles the complex evolutionary history of DNA replication in archaea. *Genome Biol Evol*, **6**, 192–212.
27. López-García,P. and Forterre,P. (1999) Control of DNA topology during thermal stress in hyperthermophilic archaea: DNA topoisomerase levels, activities and induced thermotolerance during heat and cold shock in Sulfolobus. *Mol. Microbiol.*, **33**, 766–777.
28. López-García,P., Zivanovic,Y., Deschamps,P. and Moreira,D. (2015) Bacterial gene import and mesophilic adaptation in archaea. *Nat. Rev. Microbiol.*, **13**, 447–456.
29. Boussau,B., Blanquart,S., Necsulea,A., Lartillot,N. and Gouy,M. (2008) Parallel adaptations to high temperatures in the Archaean eon. *Nature*, **456**, 942–945.
30. Groussin,M. and Gouy,M. (2011) Adaptation to environmental temperature is a major determinant of molecular evolutionary rates in archaea. *Mol. Biol. Evol.*, **28**, 2661–2674.
31. Catchpole,R.J. and Forterre,P. (2019) The evolution of reverse gyrase suggests a nonhyperthermophilic last universal common ancestor. *Mol. Biol. Evol.*, **36**, 2737–2747.
32. Kikuchi,A. and Asai,K. (1984) Reverse gyrase–a topoisomerase which introduces positive superhelical turns into DNA. *Nature*, **309**, 677–681.
33. Forterre,P., Mirambeau,G., Jaxel,C., Nadal,M. and Duguet,M. (1985) High positive supercoiling in vitro catalyzed by an ATP and polyethylene glycol-stimulated topoisomerase from *Sulfolobus acidocaldarius*. *EMBO J.*, **4**, 2123–2128.
34. Atomi,H., Matsumi,R. and Imanaka,T. (2004) Reverse gyrase is not a prerequisite for hyperthermophilic life. *J. Bacteriol.*, **186**, 4829–4833.
35. Lipscomb,G.L., Hahn,E.M., Crowley,A.T. and Adams,M.W.W. (2017) Reverse gyrase is essential for microbial growth at 95 °C. *Extremophiles*, **21**, 603–608.
36. Gadelle,D., Filée,J., Buhler,C. and Forterre,P. (2003) Phylogenomics of type II DNA topoisomerases: review articles. *Bioessays*, **25**, 232–242.
37. Santangelo,T.J., Čuboňová,L. and Reeve,J.N. (2010) *Thermococcus kodakarensis* genetics: TK1827-encoded β-glycosidase, new positive-selection protocol, and targeted and repetitive deletion technology. *Appl. Environ. Microbiol.*, **76**, 1044–1052.
38. Huber,R., Sleytr,U.B. and Stetter,K.O. (1986) *Thermotoga maritima* sp. nov. represents a new genus of unique extremely thermophilic eubacteria growing up to 90∼C. *Arch. Microbiol.*, **144**, 324–333.
39. Atomi,H., Fukui,T., Kanai,T., Morikawa,M. and Imanaka,T. (2004) Description of *Thermococcus kodakaraensis* sp. nov., a well studied hyperthermophilic archaeon previously reported as Pyrococcus sp. KOD1. *Archaea*, **1**, 263–267.
40. Guipaud,O., Marguet,E., Noll,K.M., de la Tour,C.B. and Forterre,P. (1997) Both DNA gyrase and reverse gyrase are present in the hyperthermophilic bacterium *Thermotoga maritima*. *Proc. Natl. Acad. Sci. U.S.A.*, **94**, 10606–10611.
41. Guipaud,O. and Forterre,P. (2001) DNA gyrase from Thermotoga maritima. *Meth. Enzymol.*, **334**, 162–171.
42. Tretter,E.M., Lerman,J.C. and Berger,J.M. (2010) A naturally chimeric type IIA topoisomerase in Aquifex aeolicus highlights an evolutionary path for the emergence of functional paralogs. *Proc. Natl. Acad. Sci. U.S.A.*, **107**, 22055–22059.
43. Catchpole,R., Gorlas,A., Oberto,J. and Forterre,P. (2018) A series of new *E. coli-Thermococcus* shuttle vectors compatible with previously existing vectors. *Extremophiles*, **22**, 591–598.
44. Sato,T., Fukui,T., Atomi,H. and Imanaka,T. (2005) Improved and versatile transformation system allowing multiple genetic manipulations of the hyperthermophilic archaeon Thermococcus kodakaraensis. *Appl. Environ. Microbiol.*, **71**, 3889–3899.
45. Akpan,I., Bankole,M.O. and Adesemowo,A.M. (1999) A rapid plate culture method for screening of amylase producing micro-organisms. *Biotechnol. Tech.*, **13**, 411–413.
46. Ducret,A., Quardokus,E.M. and Brun,Y.V. (2016) MicrobeJ, a tool for high throughput bacterial cell detection and quantitative analysis. *Nat Microbiol*, **1**, 16077.
47. Gorlas,A. and Geslin,C. (2013) A simple procedure to determine the infectivity and host range of viruses infecting anaerobic and hyperthermophilic microorganisms. *Extremophiles*, **17**, 349–355.
48. López-García,P. and Forterre,P. (1997) DNA topology in hyperthermophilic archaea: reference states and their variation with

growth phase, growth temperature, and temperature stresses. *Mol. Microbiol.*, **23**, 1267–1279.
49. Rueden,C.T., Schindelin,J., Hiner,M.C., DeZonia,B.E., Walter,A.E., Arena,E.T. and Eliceiri,K.W. (2017) ImageJ2: ImageJ for the next generation of scientific image data. *BMC Bioinformatics*, **18**, 529.
50. Charbonnier,F., Erauso,G., Barbeyron,T., Prieur,D. and Forterre,P. (1992) Evidence that a plasmid from a hyperthermophilic archaebacterium is relaxed at physiological temperatures. *J. Bacteriol.*, **174**, 6103–6108.
51. Li,H. and Durbin,R. (2009) Fast and accurate short read alignment with Burrows–Wheeler transform. *Bioinformatics*, **25**, 1754–1760.
52. Love,M.I., Huber,W. and Anders,S. (2014) Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome Biol.*, **15**, 550.
53. Tatusov,R.L., Galperin,M.Y., Natale,D.A. and Koonin,E.V. (2000) The COG database: a tool for genome-scale analysis of protein functions and evolution. *Nucleic Acids Res.*, **28**, 33–36.
54. Horowitz,D.S. and Wang,J.C. (1987) Mapping the active site tyrosine of *Escherichia coli* DNA gyrase. *J. Biol. Chem.*, **262**, 5339–5344.
55. Critchlow,S.E. and Maxwell,A. (1996) DNA cleavage is not required for the binding of quinolone drugs to the DNA Gyrase−DNA complex. *Biochemistry*, **35**, 7387–7393.
56. Gibson,E.G., Oviatt,A.A. and Osheroff,N. (2020) Two-Dimensional gel electrophoresis to resolve DNA topoisomers. In: Hanada,K. (ed). *DNA Electrophoresis: Methods and Protocols, Methods in Molecular Biology*. Springer US, NY, pp. 15–24.
57. Gorlas,A., Catchpole,R., Marguet,E. and Forterre,P. (2019) Increase of positive supercoiling in a hyperthermophilic archaeon after UV irradiation. *Extremophiles*, **23**, 141–149.
58. López-García,P., Forterre,P., van der Oost,J. and Erauso,G. (2000) Plasmid pGS5 from the hyperthermophilic archaeon archaeoglobus profundus is negatively supercoiled. *J. Bacteriol.*, **182**, 4998–5000.
59. Mustaev,A., Malik,M., Zhao,X., Kurepina,N., Luan,G., Oppegard,L.M., Hiasa,H., Marks,K.R., Kerns,R.J., Berger,J.M. *et al.* (2014) Fluoroquinolone-gyrase-DNA complexes. *J. Biol. Chem.*, **289**, 12300–12312.
60. Gellert,M., Mizuuchi,K., O'Dea,M.H., Itoh,T. and Tomizawa,J.-I. (1977) Nalidixic acid resistance: A second genetic character involved in DNA gyrase activity. *Proc. Natl. Acad. Sci. U.S.A.*, **74**, 4772–4776.
61. Sugino,A., Peebles,C.L., Kreuzer,K.N. and Cozzarelli,N.R. (1977) Mechanism of action of nalidixic acid: Purification of Escherichia coli nalA gene product and its relationship to DNA gyrase and a novel nicking-closing enzyme. *Proc. Natl Acad. Sci. U.S.A.*, **74**, 4767–4771.
62. Pan,L., Li,J., Li,C., Tang,X., Yu,G. and Wang,Y. (2018) Study of ciprofloxacin biodegradation by a Thermus sp. isolated from pharmaceutical sludge. *J. Hazard. Mater.*, **343**, 59–67.
63. Bell,S.D., Jaxel,C., Nadal,M., Kosa,P.F. and Jackson,S.P. (1998) Temperature, template topology, and factor requirements of archaeal transcription. *Proc. Natl. Acad. Sci. U.S.A.*, **95**, 15218–15222.
64. El Houdaigui,B., Forquet,R., Hindré,T., Schneider,D., Nasser,W., Reverchon,S. and Meyer,S. (2019) Bacterial genome architecture shapes global transcriptional regulation by DNA supercoiling. *Nucleic Acids Res.*, **47**, 5648–5657.
65. Dorman,C.J. (2019) DNA supercoiling and transcription in bacteria: a two-way street. *BMC Mol Cell Biol*, **20**, 26.
66. Fujikane,R., Ishino,S., Ishino,Y. and Forterre,P. (2010) Genetic analysis of DNA repair in the hyperthermophilic archaeon, *Thermococcus kodakaraensis*. *Genes Genet. Syst.*, **85**, 243–257.
67. Takahashi,T.S., Da Cunha,V., Krupovic,M., Mayer,C., Forterre,P. and Gadelle,D. (2020) Expanding the type IIB DNA topoisomerase family: identification of new topoisomerase and topoisomerase-like proteins in mobile genetic elements. *NAR Genomics Bioinformatics*, **2**, lqz021.
68. Sanders,T.J., Ullah,F., Gehring,A.M., Burkhart,B.W., Vickerman,R.L., Fernando,S., Gardner,A.F., Ben-Hur,A. and Santangelo,T.J. (2021) Extended archaeal histone-based chromatin structure regulates global gene expression in *Thermococcus kodakarensis*. *Front. Microbiol.*, **12**, 681150.
69. Jäger,D., Förstner,K.U., Sharma,C.M., Santangelo,T.J. and Reeve,J.N. (2014) Primary transcriptome map of the hyperthermophilic archaeon *Thermococcus kodakarensis*. *BMC Genomics*, **15**, 684.
70. Bailey,T.L., Johnson,J., Grant,C.E. and Noble,W.S. (2015) The MEME Suite. *Nucleic Acids Res.*, **43**, W39–W49.
71. Gadelle,D., Bocs,C., Graille,M. and Forterre,P. (2005) Inhibition of archaeal growth and DNA topoisomerase VI activities by the Hsp90 inhibitor radicicol. *Nucleic Acids Res.*, **33**, 2310–2317.
72. Lindahl,T. (1993) Instability and decay of the primary structure of DNA. *Nature*, **362**, 709–715.
73. Hethke,C., Bergerat,A., Hausner,W., Forterre,P. and Thomm,M. (1999) Cell-dree transcription at 95°: thermostability of transcriptional components and DNA topology requirements of pyrococcus transcription. *Genetics*, **152**, 1325–1333.
74. Ferrándiz,M.-J., Martín-Galiano,A.J., Arnanz,C., Camacho-Soguero,I., Tirado-Vélez,J.-M. and de la Campa,A.G. (2016) An increase in negative supercoiling in bacteria reveals topology-reacting gene clusters and a homeostatic response mediated by the DNA topoisomerase I gene. *Nucleic Acids Res.*, **44**, 7292–7303.
75. Pruss,G.J., Manes,S.H. and Drlica,K. (1982) Escherichia coli DNA topoisomerase I mutants: Increased supercoiling is corrected by mutations near gyrase genes. *Cell*, **31**, 35–42.
76. Sandman,K. and Reeve,J.N. (2006) Archaeal histones and the origin of the histone fold. *Curr. Opin. Microbiol.*, **9**, 520–525.
77. Peeters,E., Driessen,R.P.C., Werner,F. and Dame,R.T. (2015) The interplay between nucleoid organization and transcription in archaeal genomes. *Nat. Rev. Microbiol.*, **13**, 333–341.
78. Maruyama,H., Harwood,J.C., Moore,K.M., Paszkiewicz,K., Durley,S.C., Fukushima,H., Atomi,H., Takeyasu,K. and Kent,N.A. (2013) An alternative beads-on-a-string chromatin architecture in *Thermococcus kodakarensis*. *EMBO Rep.*, **14**, 711–717.
79. Nalabothula,N., Xi,L., Bhattacharyya,S., Widom,J., Wang,J.-P., Reeve,J.N., Santangelo,T.J. and Fondufe-Mittendorf,Y.N. (2013) Archaeal nucleosome positioning in vivo and in vitro is directed by primary sequence motifs. *BMC Genomics*, **14**, 391.
80. Sanders,T.J., Marshall,C.J. and Santangelo,T.J. (2019) The role of archaeal chromatin in transcription. *J. Mol. Biol.*, **431**, 4103–4115.
81. Bailey,K.A., Chow,C.S. and Reeve,J.N. (1999) Histone stoichiometry and DNA circularization in archaeal nucleosomes. *Nucleic Acids Res.*, **27**, 532–536.
82. Yunwei,X. and Reeve John,N. (2004) Transcription by an Archaeal RNA polymerase is slowed but not blocked by an archaeal nucleosome. *J. Bacteriol.*, **186**, 3492–3498.
83. Mattiroli,F., Bhattacharyya,S., Dyer,P.N., White,A.E., Sandman,K., Burkhart,B.W., Byrne,K.R., Lee,T., Ahn,N.G., Santangelo,T.J. *et al.* (2017) Structure of histone-based chromatin in Archaea. *Science*, **357**, 609–612.
84. Rojec,M., Hocher,A., Stevens,K.M., Merkenschlager,M. and Warnecke,T. (2019) Chromatinization of *Escherichia coli* with archaeal histones. *Elife*, **8**, e49038.
85. Sioud,M., Possot,O., Elie,C., Sibold,L. and Forterre,P. (1988) Coumarin and quinolone action in archaebacteria: evidence for the presence of a DNA gyrase-like enzyme. *J. Bacteriol.*, **170**, 946–953.
86. Sioud,M., Baldacci,G., de Recondo,A.M. and Forterre,P. (1988) Novobiocin induces positive supercoilling of small plasmids from halophilic arctiaebacterla in vivo. *Nucleic Acids Res.*, **16**, 1379–1391.
87. Holmes,M.L. and Dyall-Smith,M.L. (1991) Mutations in DNA gyrase result in novobiocin resistance in halophilic archaebacteria. *J. Bacteriol.*, **173**, 642–648.
88. Yang,C.F., Kim,J.M., Molinari,E. and DasSarma,S. (1996) Genetic and topological analyses of the bop promoter of *Halobacterium halobium*: stimulation by DNA supercoiling and non-B-DNA structure. *J. Bacteriol.*, **178**, 840–845.
89. Tarasov,V., Schwaiger,R., Furtwängler,K., Dyall-Smith,M. and Oesterhelt,D. (2011) A small basic protein from the brz-brb operon is involved in regulation of bop transcription in *Halobacterium salinarum*. *BMC Mol. Biol.*, **12**, 42.
90. Bizard,A. (2011) Caractérisation des topoisomérases de type IA de l'archaeon hyperthermophile *Sulfolobus solfataricus*.
91. Couturier,M., Gadelle,D., Forterre,P., Nadal,M. and Garnier,F. (2020) The reverse gyrase TopR1 is responsible for the homeostatic control of DNA supercoiling in the hyperthermophilic archaeon *Sulfolobus solfataricus*. *Mol. Microbiol.*, **113**, 356–368.
92. Kampmann,M. and Stock,D. (2004) Reverse gyrase has heat-protective DNA chaperone activity independent of supercoiling. *Nucleic Acids Res.*, **32**, 3537–3545.
93. Napoli,A., Valenti,A., Salerno,V., Nadal,M., Garnier,F., Rossi,M. and Ciaramella,M. (2004) Reverse gyrase recruitment to DNA after

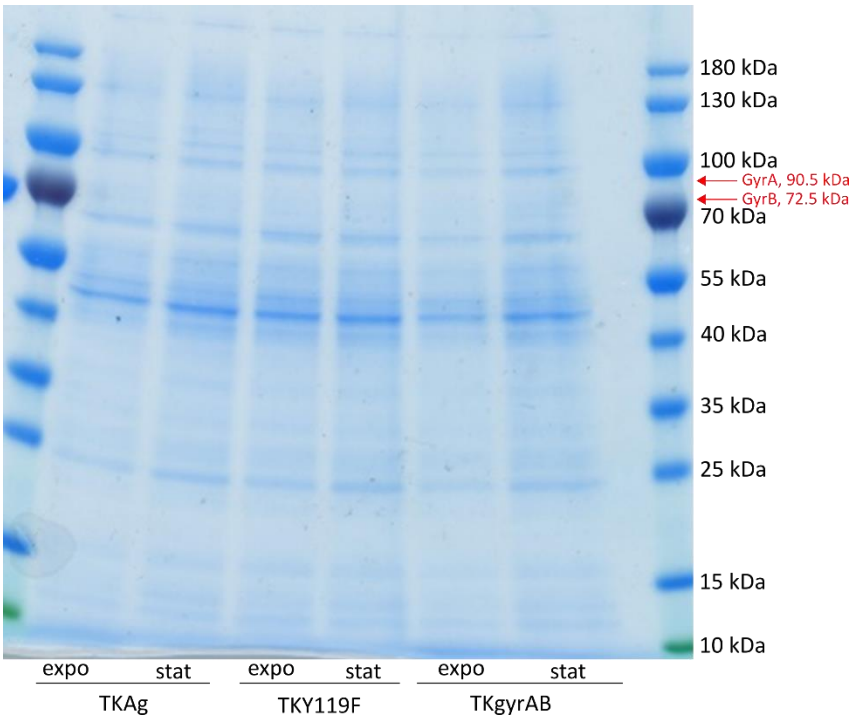
UV light irradiation in Sulfolobus solfataricus. *J. Biol. Chem.*, **279**, 33192–33198.
94. Hsieh,T. and Plank,J.L. (2006) Reverse gyrase functions as a DNA renaturase: annealing of complementary single-stranded circles and positive supercoiling of a bubble substrate *. *J. Biol. Chem.*, **281**, 5640–5647.
95. Valenti,A., Napoli,A., Ferrara,M.C., Nadal,M., Rossi,M. and Ciaramella,M. (2006) Selective degradation of reverse gyrase and DNA fragmentation induced by alkylating agent in the archaeon *Sulfolobus solfataricus*. *Nucleic Acids Res.*, **34**, 2098–2108.
96. Forterre,P. (2002) A hot story from comparative genomics: reverse gyrase is the only hyperthermophile-specific protein. *Trends Genet.*, **18**, 236–237.
97. Meyer,S., Jost,D., Theodorakopoulos,N., Peyrard,M., Lavery,R. and Everaers,R. (2013) Temperature dependence of the DNA double helix at the nanoscale: Structure, elasticity, and fluctuations. *Biophys. J.*, **105**, 1904–1914.
98. Marguet,E., Zivanovic,Y. and Forterre,P. (1996) DNA topological change in the hyperthermophilic archaeon *Pyrococcus abyssi* exposed to low temperature. *FEMS Microbiol. Lett.*, **142**, 31–36.
99. López-García,P. (1999) DNA supercoiling and temperature adaptation: a clue to early diversification of life? *J. Mol. Evol.*, **49**, 439–452.

SUPPLEMENTARY FIGURES AND TABLES



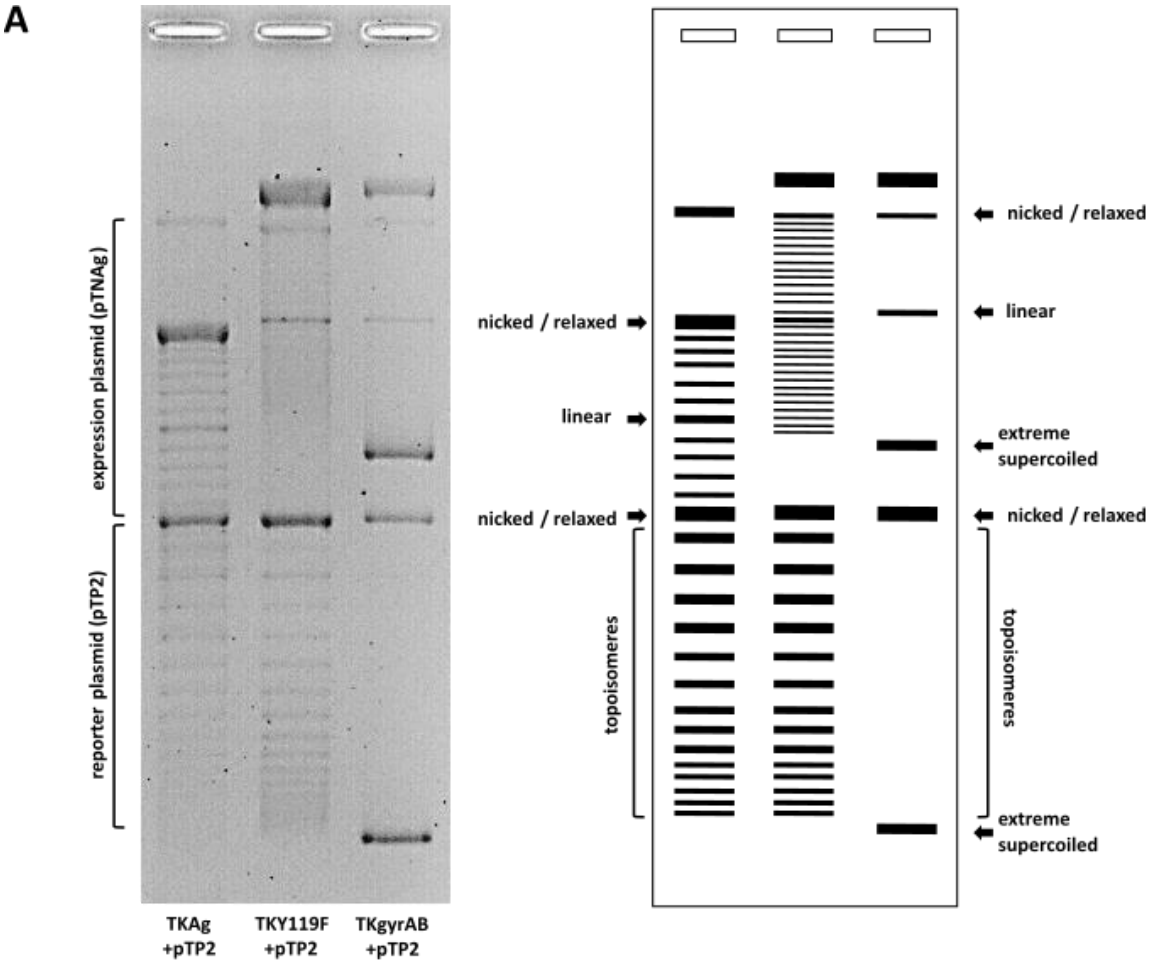
Supplementary figure 1. The *gyrA* and *gyrB* from *T. maritima* are transcribed in *T. kodakarensis* TKgyrAB

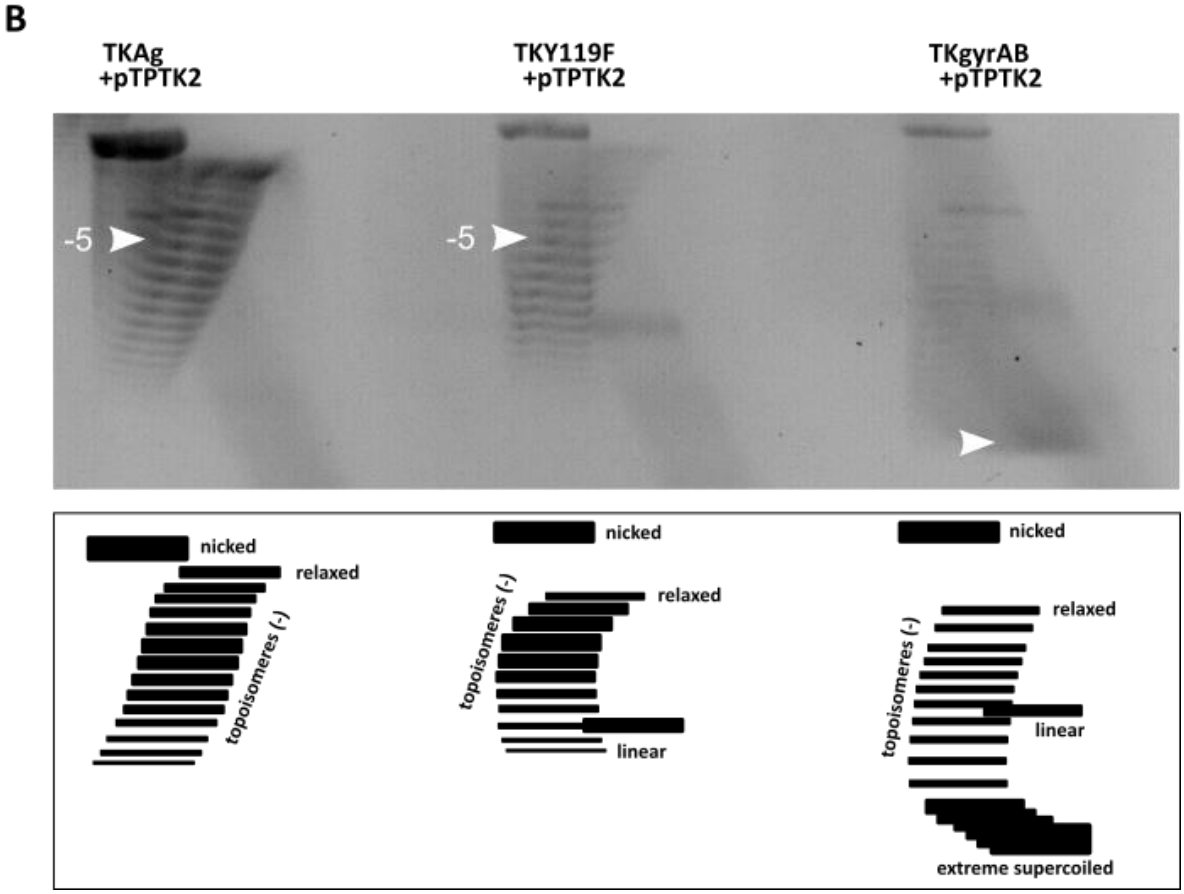
The expression of *gyrA* and *gyrB* was analysed by RT-PCR for three independent clones of strain TKgyrAB. For each clone, total RNA was isolated and then retrotranscribed to cDNA (see material and methods). The total RNA (left side of the gels) or cDNA (right side of the gels) was used as template for PCR and the obtained products were separated by agarose gel electrophoresis. The position of the used specific oligonucleotides is indicated with arrows. T is a control PCR where total RNA or cDNA from *T. kodakarensis* TKAg which carries the empty expression vector was used as template. TK0149 is plasmid-encoded pyruvoyl-dependent arginine decarboxylase which confers prototrophy to agmatine.



Supplementary figure 2. SDS-PAGE analysis of soluble proteins from *T. kodakarensis*.

TKAg, TKY119F and TKgyrAB were grown until exponential (expo) or stationary phase (stat). Cells were pelleted and resuspended in Laemmli buffer and heat-denatured. Resulting cell lysates were analysed on a 4-20% polyacrylamide gel. Protein ladder sizes are indicated on the right side of the gel. Expected bands at 90,5 kDa and 72,5 are not visible in TKY119F and TKgyrAB lysates indicating that *gyrA* and *gyrB* are not overexpressed.

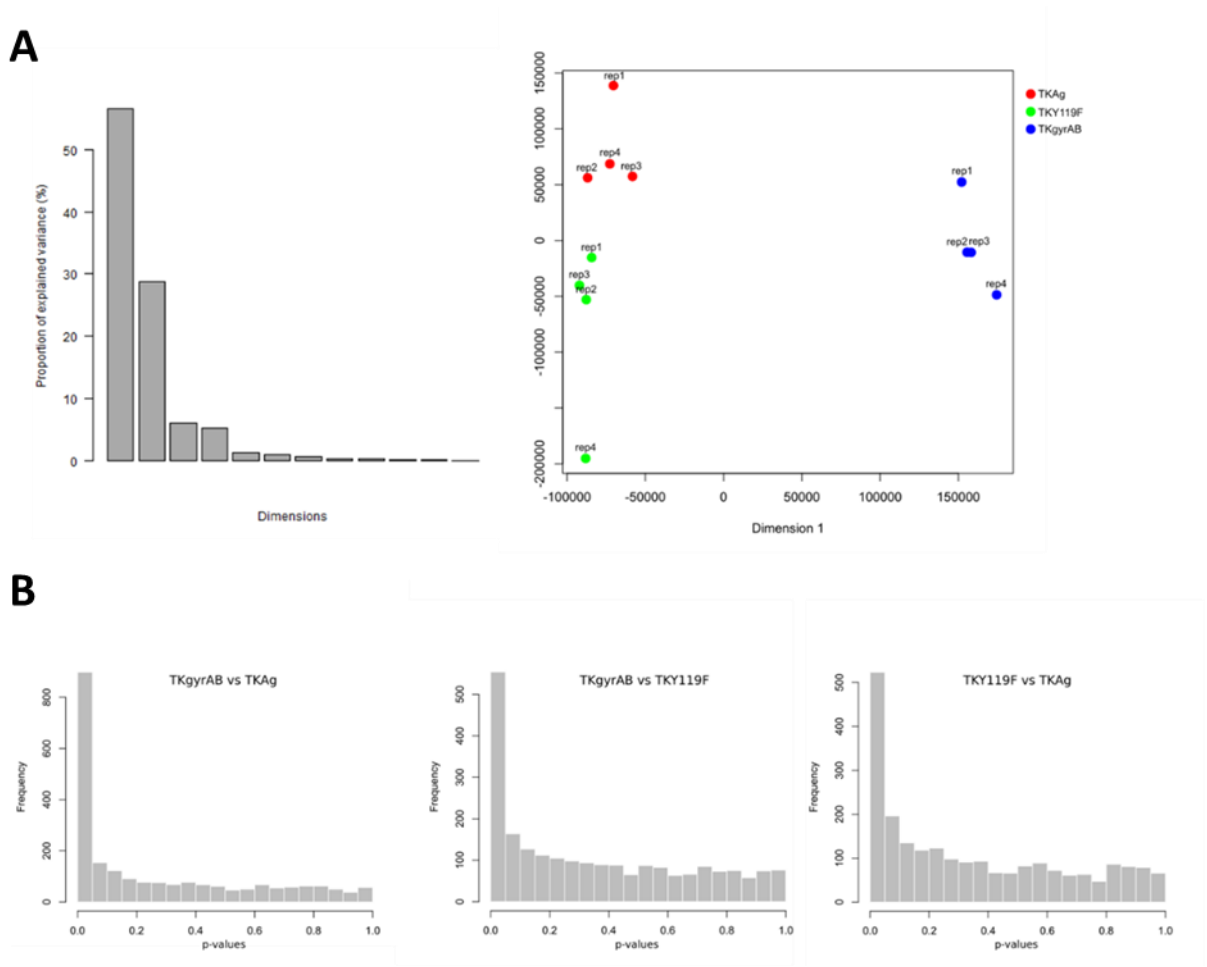

A
expression plasmid (pTNAg)
reporter plasmid (pTP2)
TKAg +pTP2
TKY119F +pTP2
TKgyrAB +pTP2
nicked / relaxed
linear
nicked / relaxed
topoisomeres
nicked / relaxed
linear
extreme supercoiled
nicked / relaxed
topoisomeres
extreme supercoiled


B
TKAg +pTPTK2
TKY119F +pTPTK2
TKgyrAB +pTPTK2
-5
-5
nicked
relaxed
topoisomeres (-)
nicked
topoisomeres (-)
relaxed
linear
nicked
relaxed
topoisomeres (-)
linear
extreme supercoiled

Supplementary figure 3. Topological profile of plasmids isolated from *Thermococcus kodakarensis* TKgyrAB, TKAg or TKY119F strains.

A) The plasmid DNA was migrated in 0.8% (w/v) agarose gel in absence of intercalating agents for 24h at 1,6 V/cm at 4°C. The ethidium bromide gel is shown on the left and its schematic representation on the right. Different topological forms of plasmids are indicated.

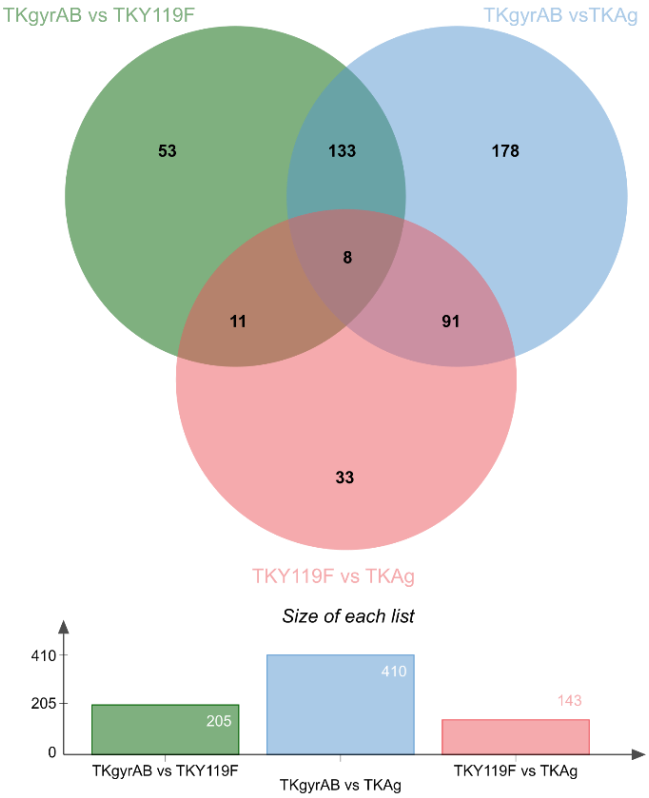
B) The plasmid DNA was migrated in two dimensions in a 1% (w/v) agarose gel prepared with 1X TEP (36 mM Tris-HCl, 30 mM NaH_2PO_4, 1 mM EDTA, pH 7.8). Chloroquine was added at 10 µg/ml only in the second dimension. First dimension was run for 15 h at 1.2 V/cm and second dimension was run for 5 h at the same voltage. Observed topological forms are indicated in the corresponding cartoon. For the sake of clarity only the part of the gel containing the reporter plasmid pTPTK2 is shown. The major topoisomer in TKAg and TKY119F strains is indicated with white arrow.



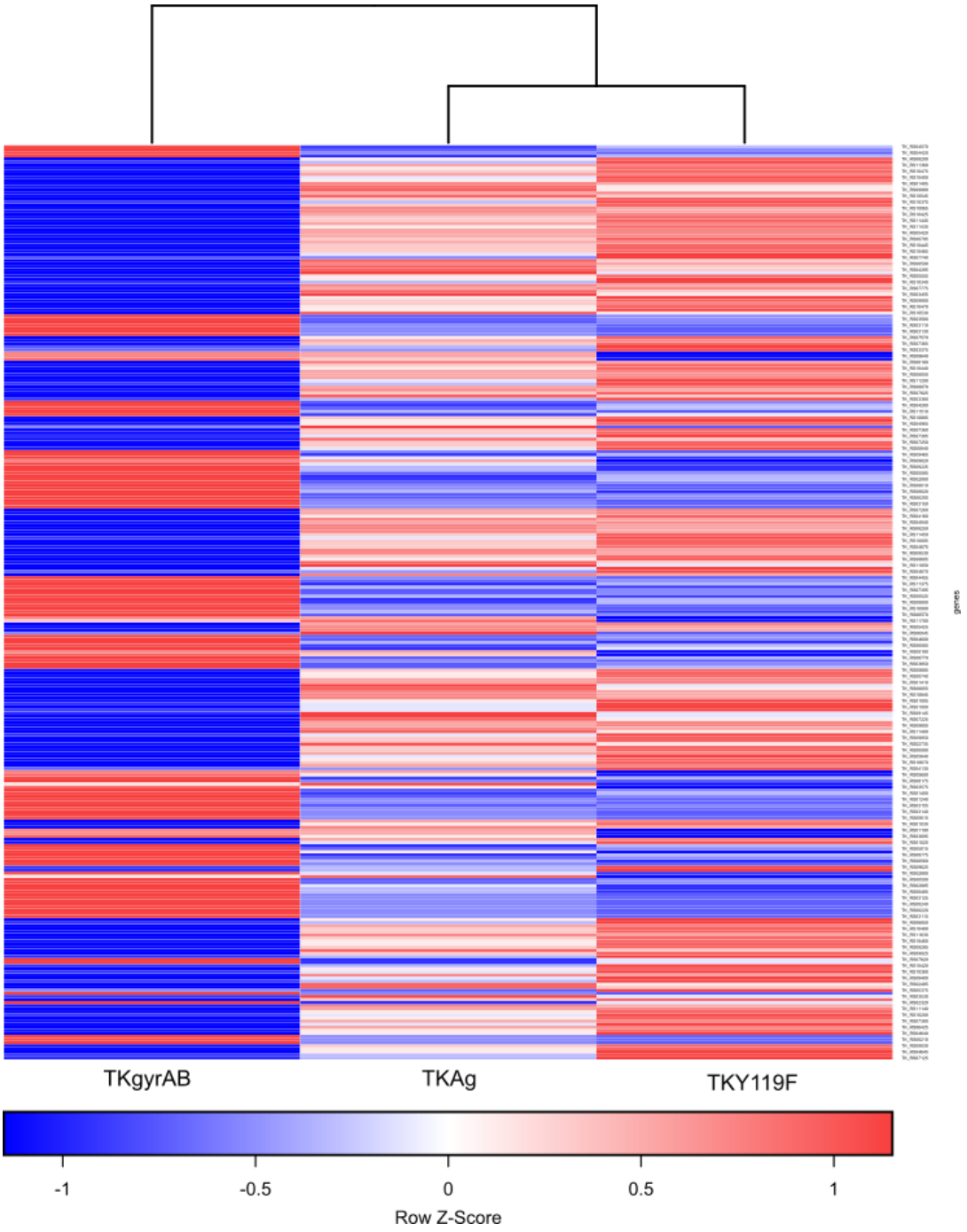
Supplementary figure 4: Quality assessment of RNA-seq data

A) Principal component analysis was performed to assess the variability in the dataset. Left graph shows the distribution of components as a function of data variability. The first two principal components capture more than 80% of the variance. In the graph on the right, each point corresponds to a single RNA-seq dataset. For each condition (strain) four biological replicates were analysed.

B) P-value histogram for the three differential analyses to evaluate the p-value significance threshold. The three plots show anti-conservative p-values distribution with the null p-values uniformly distributed between 0 and 1. Such distribution indicates that the frequency of the false positives in the three datasets will be low for the p-values less than 0.05.

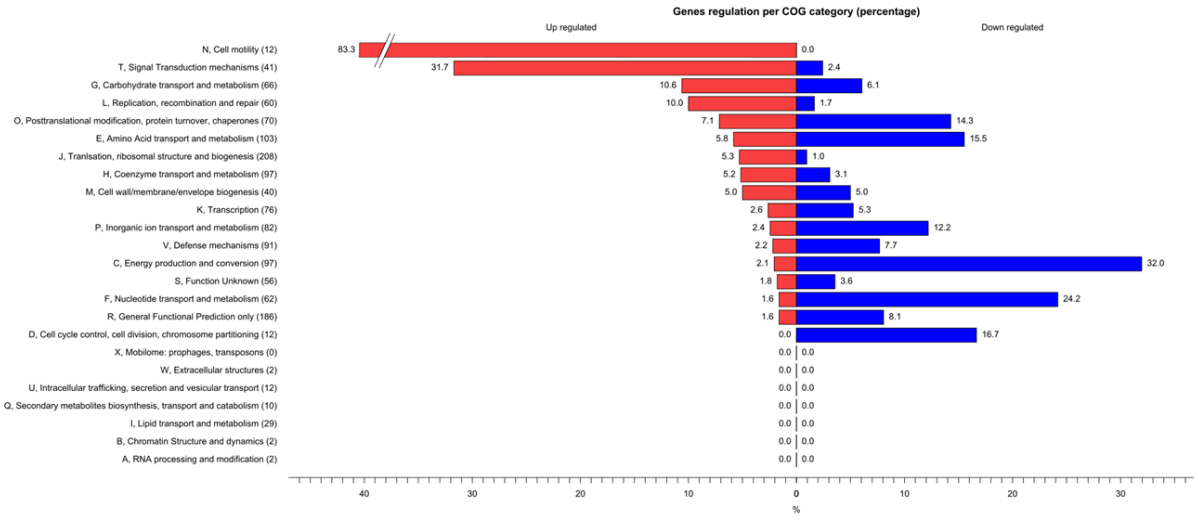


Supplementary figure 5: Venn diagram showing the overlap between the DEGs detected in three pairwise comparisons of transcriptomes. Green circle: genes that react to negative supercoiling; Pink circle : genes that react to gyrase expression and DNA binding; Blue circle : genes that react to negative supercoiling, gyrase expression and DNA binding. While there is significant overlap between pink and blue as well as green and blue circles, only few genes are shared between green and pink circle. This indicates that the genes identified by TKGyrAB vs TKY119F comparison are enriched in genes responding specifically to negative supercoiling activity of the gyrase.



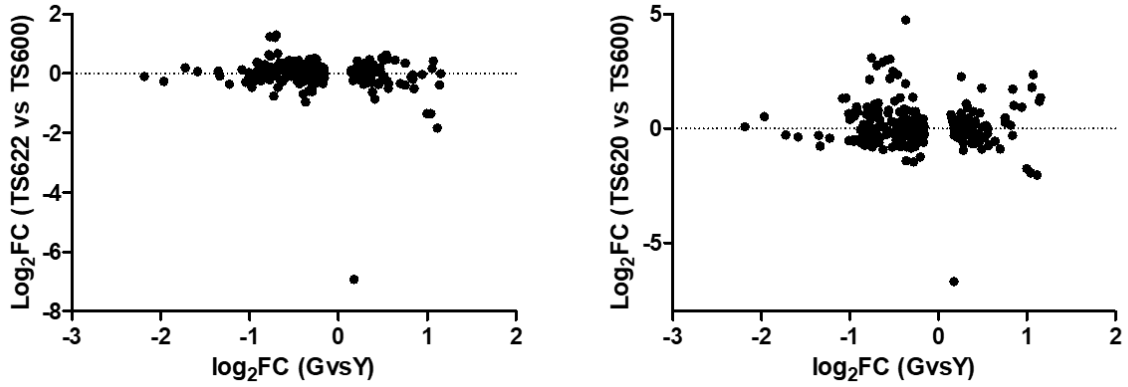
Supplementary figure 6. Relative count abundancies for supercoiling – responding subset of DEGs using Z-score scaling.

Each line corresponds to one gene and each column correspond to the one of the three recombinant strains as indicated below the heat map. The Z score for each gene is the mean value of four individual Z-scores obtained for each biological replicate. Individual Z-scores were calculated using normalised counts for each gene according to the standard equation. The two control strains (TKAg and TKY119F) are clustered together to exclusion of the gyrase expressing strain (TKGyrAB).



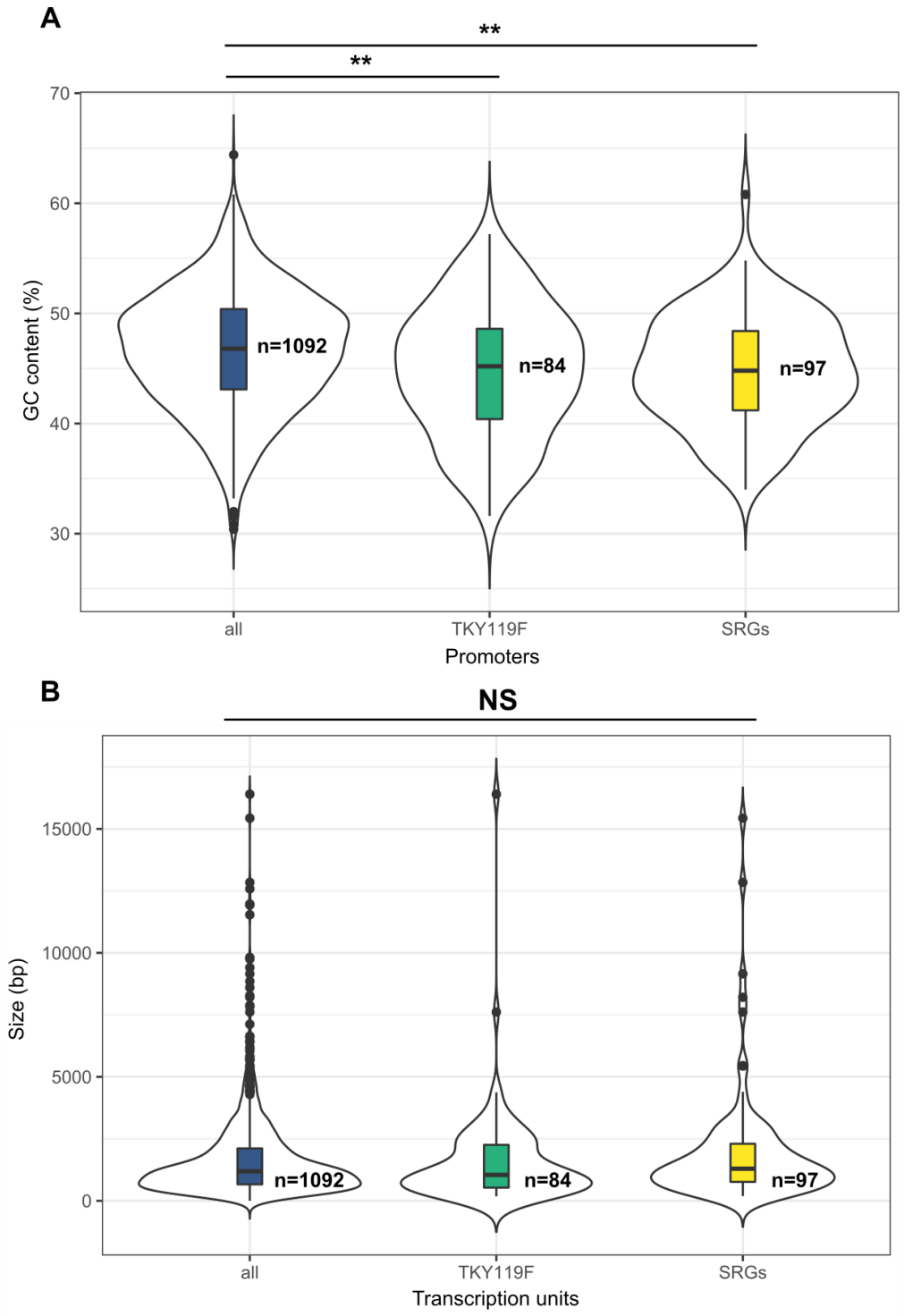
Supplementary figure 7: Assignment of DEGs to ArCOG functional categories.

Total number of genes assigned to each category is indicated in the brackets. The percentage of DEGs in each category is indicated next to each bar. Note that 1416 out of 2256 annotated protein coding genes of *T. kodakarensis* TS559 have been assigned functions.



Supplementary figure 8: Correlation analysis of transcriptional output for genes responding both to supercoiling and chromatin defect.

Each dot corresponds to log_2FC values for SRGs with Padj < 0.05. The top 30 upregulated SRGs were removed from the analysis. Graph on the left shows the correlation with corresponding genes from *T. kodakarensis* TS622 ($HTkB^{WT}$ $HTkA^{G17D}$) and on the left the correlation with TS620 (ΔHTkB $HTkA^{G17D}$). Non-parametric Spearman correlation test revealed the absence of significant correlation between the two sets of data.

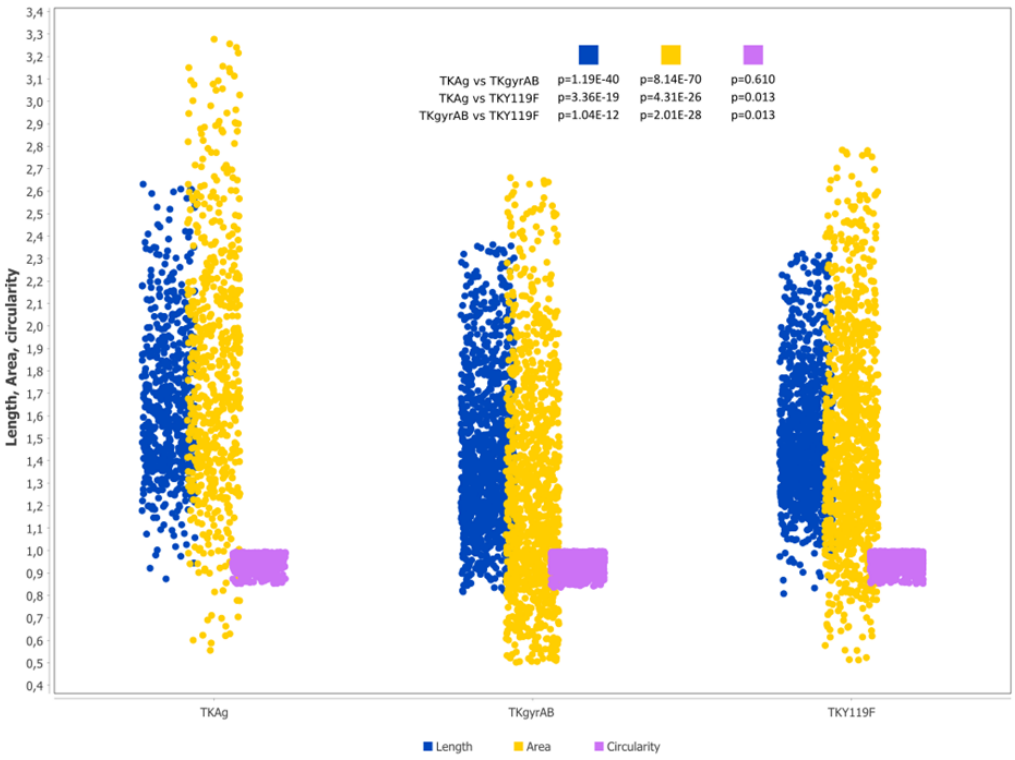


Supplementary figure 9: Analysis of GC content of SRGs promotor regions and length of the corresponding transcripts

Operons were predicted for *T. kodakarensis* TS559 genome using Operon-mapper software (1). The 250 bp upstream of each operon start and the predicted transcription unit size were extracted using R and the data were plotted using ggplot2 package.

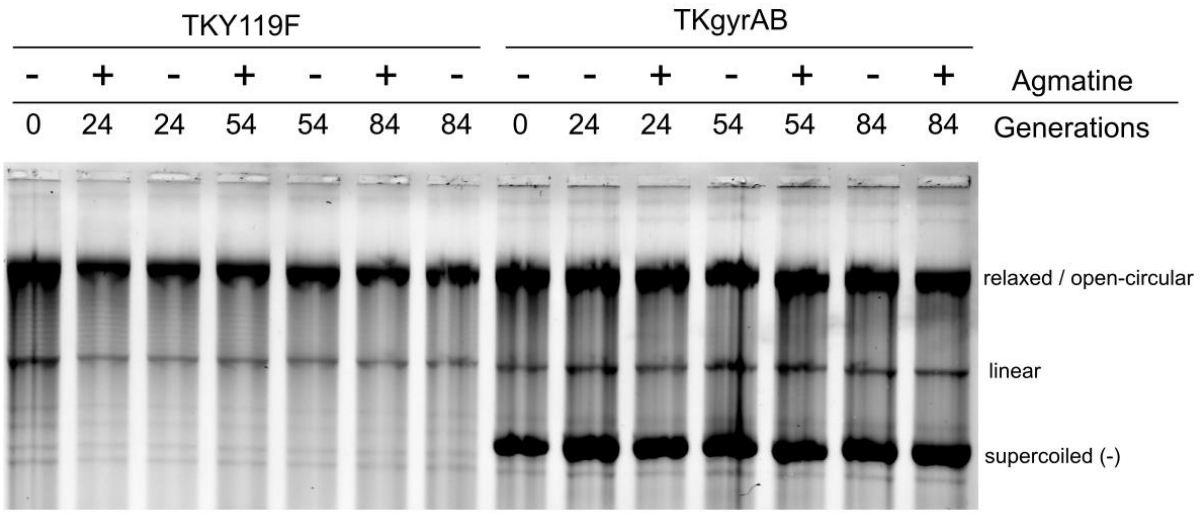
A) GC content of promoter regions of SRGs (SRGs) were compared to all predicted promoter regions (all) in *T. kodakarensis* TS559 and to the promoter regions of the genes reacting to expression of the catalytic mutant (TKY119F). The median values were compared using Mann and Whitney non-parametric statistical test. Significantly different medians are indicated by ** (p value < 0.01) on the top of the graph.

(B) The median values of transcription units size analysis were compared by Kruskal-Wallis non-parametric statistical test. None of the medians tested is significantly different from the others (p value > 0.05).



Supplementary figure 10. Analysis of cell size and shape of recombinant *T. kodakaresis* cells by DIC microscopy.

Box plots showing the distribution of the cell surface for the indicated number of cells as determined using MicrobeJ. The mean values were compared using nonparametric statistical test (Mann-Whitney U test) and the statistical significance is expressed as P values. Both TKY119F and TKGyrAB strains are significantly smaller than the control TKAg strain but the shape of TKGyrAB is not significantly different from the control strain.



Supplementary figure 11. One dimensional gel agarose analysis of pTPTK2 topology over prolonged subculturing of *Thermococcus kodakarensis*.

pTPTK2 plasmids were isolated from the cultures after 24, 54 or 84 generations and separated on the 0.8% agarose gel in absence of chloroquine. The presence (non-selective condition) or absence (selective condition) of agmatine in culture medium is indicated above the gel.

Supplementary Table 1: Strains and plasmids used in this work

Strain name	Genotype	Markers	Source
Escherichia coli XL1-Blue	*endA*1 *gyrA*96 thi-1 *recA*1 *relA*1 lac *glnV*44 F'[::Tn10 *proAB*$^{+}$ *lacI*^{q} Δ(l*acZ*)M15] *hsdR*17	Tetracycline resistance Nalidixic acid resistance	Stratagene
Thermococcus kodakarensis TS559	Δ*pyrF*; Δ*trpE::pyrF*, Δ*TK0664*, Δ*TK0149*	Uracil prototrophy Tryptophan auxotrophy Agmatine auxotrophy 6-methylpurine resistant	Santangelo 2010

Plasmid name	Genotype	*E. coli* marker(s)	*T. kodakarenesis* marker(s)	Source (Accession No.)
pTNAg	pLC70Δ(*TK0254-PF1848*):: (P_{TK0149}-*TK0149*)	AmpR, KanR	Agm	Catchpole 2018 (MG920813)
pLC70	see reference	AmpR, KanR	Trp, MevR	Santangelo 2008 (N/A)
pTPTK2	pTP2::(p15A-*cat*),(P_{TK2279}-*TK0254*)	CmR	Trp	Catchpole 2018 (MG920816)
pTNAg-gyrAB	pLC70Δ(*TK0254-PF1848*):: (P_{TK0149}-*TK0149*)	AmpR, KanR	Agm	*this work*
pTNAg-Y119F	pLC70Δ(*TK0254-PF1848*):: (P_{TK0149}-*TK0149*)	AmpR, KanR	Agm	*this work*

AmpR=ampicillin resistance; KanR=kanamycin resistance; CmR=chloramphenicol resistance; Trp=tryptophan prototrophy in a Δ*trpE* (TK0254) background; MevR=mevinolin resistance; Agm=agmatine prototrophy in a ΔTK0149 background.

Supplementary Table 2: Oligonucleotides used in this work

Primer name	Template sequence	Sequence (5′-3′)
Tmar.gyrA.1	*T. maritima MBS8* genome	gtgagcaaaatgcgcgttgcggcagtcgacagcgatatatttatatagggatatagtaatagataatatcacaggtggtatgaATGCCAGAGATCCTGATAAAC
Tmar.gyrA.2	*T. maritima MBS8* genome	tccattcataccacctgGGTATGTTCTATGGGTTTCC
Tmar.gyrB.1	*T. maritima MBS8* genome	catagaacatacccaggtggtatgaATGGAAAAGTACTCCGCTG
Tmar.gyrB.2	*T. maritima MBS8* genome	acgttcatcaaagttcatctagagcggccgCTAGATATCCAGTTCTTTCACTTTC
pTNAg.GA.gyrAB.1	pLC70-gyrAB	tacgccaagcttggtaccgagctcgTCGACAGCGATATATTTATATAGG
pTNAg.GA.gyrAB.2	pLC70-gyrAB	tccattcataccacctGGGTATGTTCTATGGGTTTCC
pTNAg.GA.Y119F.1	pTNAg-gyrAB	tacgccaagcttggtaccgagctcgGTCGACAGCGATATATTTATATAGGG
pTNAg.GA.Y119F.2	pTNAg-gyrAB	gcctcgtgagtctCGCTTCCGTGAACCTCATC
pTNAg.GA.Y119F.3	pTNAg-gyrAB	gttcacggaagcgAGACTCACGAGGCTCGCAG
pTNAg.GA.Y119F.4	pTNAg-gyrAB	agcacactggcggccgttactagtgGGCTAGATATCCAGTTCTTTCACTTTC
RT-PCR.gyrA.1	pTNAg-gyrAB / pTNAg-Y119F	GTCGCGAAGAACACCTCATC
RT-PCR.gyrA.2	pTNAg-gyrAB / pTNAg-Y119F	TTGCCGAATCCCTTCTCTGT
RT-PCR.gyrB.1	pTNAg-gyrAB / pTNAg-Y119F	GCAAAACAGGCCAGAGACAG
RT-PCR.gyrB.2	pTNAg-gyrAB / pTNAg-Y119F	CACTTTCAGAGCGTGCCTTT
RT-PCR.gyrAB.1	pTNAg-gyrAB / pTNAg-Y119F	AGGGATTCGGCAAGAGAACA
RT-PCR.gyrAB.2	pTNAg-gyrAB / pTNAg-Y119F	TCCTCGACTTCCACACTTCC

Uppercase indicates identity to the template sequence ; lowercase indicates primer extension for Gibson assembly.

Primers Tmar.gyrA.1, Tmar.gyrA.2, Tmar.gyrB.1, Tmar.gyrB.2 were used to construct pLC70-gyrAB plasmid (unpublished data). pLC70-gyrAB was used as a template to construct pTNAg-gyrAB and pTNAg-Y119F.

Supplementary Table 3: Superhelical densities from various plasmids and viruses

Organism	Optimal growth temperature	Plasmid	Superhelical density	References
Escherichia coli	37	pTZ18	-0.051	Charbonnier F. and Forterre P., *J Bacteriol.* 1994
Escherichia coli	37	pBR322	-0.050	
Escherichia coli	37	M13mp19	-0.049	
Thermus sp. YS45	65	pTYS45-1	-0.057	
Rhodothermus marinus R21	65	pRM21	-0.063	
Thermus thermophilus HB8	80	pTT8	-0.059	
Halobacterium halobium GRB	37	pGRB	-0.068	
Haloferax volcanii WR11	37	pHV11	-0.068	
Haloferax volcanii WR12	37	pHV12	-0.060	
Halobacterium volcanii DS2	37	pHV2	-0.060	
Methanococcus sp. C5	30	pURB500	-0.058	
Methanosarcina acetivorans C2A	35	pC2A	-0.048	
Methanobacterium thermoautotrophicum Marburg DSM 2133	65	pME2001	-0.013	
Sulfolobus shibatae DSM 5389	78	pSSV1	+0.015	
Desulfurolobus ambivalens DSM 3772	80	pSL10	+0.007	
Pyrococcus abyssi GE5	95	pGT5	-0.003	
Thermococcus sp. GE31	80	pGN31	+0.035	López-García P. and Forterre P., *Mol. Microbiol.* 1997
Sulfolobus islandicus REN1H1	80	pRN1	+0.008	
Sulfolobus islandicus REN1H1	80	pRN2	+0.008	
Sulfolobus sp. NZ 54/3	80	pTAU4	+0.017	
Archaeoglobus profundus	80	pGS5	-0.033	Lopez-Garcia P. et al., *J Bacteriol*. 2000
SV40 virion	37 (host)	/	-0.051	Shure M. et al., *Nucleic Acids Res.* 1977
intracellular SV40		/	-0.050	
polyoma virion		/	-0.053	
intracellular polyoma		/	-0.050	
PM2 (marine bacteriophage)	37 (host)	/	-0.017	

Chlamydia trachomatis	37	pCT-L2	-0.07	Niehus E. et al., *J Bacteriol.* 2008
mammalian COS cells	37	pRSSVO	-0.0581	Tong W. et al., *J Mol Biol.* 2006
mammalian COS cells	37	pTEKO	-0.0583	
mammalian COS cells	37	pOS47	-0.0589	
mammalian COS cells	37	pOS67	-0.0556	
Saccharomyces cerevisiae	30?	pRSSVO	-0.0421	
Saccharomyces cerevisiae	30	TAC	-0.0511	Shen CH. et al., *Mol Cell Biol.* 2001
Saccharomyces cerevisiae	30	TA-HIS3	-0.0453	Kim Y, Clark DJ., *Proc Natl Acad Sci U S A.* 2002
Saccharomyces cerevisiae	30	TRP1ARS1	-0.0506	Pederson DS et al., *Proc Natl Acad Sci U S A.* 1986
Bacillus subtilis	37	pUB110	-0.074	calculated from: Nicholson WL, Setlow P., *J Bacteriol*. 1990
Mycobacterium tuberculosis	37	pSUM36	-0.060	García MT et al., *Front Microbiol*. 2018
Thermococcus nautili	87.5	pTN1	-0.002	Gorlas A. et al., *Extremophiles*. 2019
Sulfolobus sp. NZ 59/2	80	pSTHA	+0.006	Charbonnier F. and Forterre P., *J Bacteriol*. 1994
Thermotoga sp. RQ7	80	pRQ7	-0.067	Guipaud O. et al., *Proc Natl Acad Sci U S A.* 1997
Streptococcus pneumoniae TBB1	37	pLS1	-0.059	Ferrandiz et al., *Nucleic Acids Research.* 2016
Streptococcus pneumoniae TBB1 (SCN treated)	37	pLS1	-0.083	

Supplementary Table 4: Differential expression profile of pTNAg encoded genes

		TKgyrAB vs TKY119				**TKY119F vs TKAg**				**TKgyrAB vs TKAg**			
Locus name	Annotation	log2FC	Padj	NC (G)	NC(Y)	log2FC	Padj	NC (Y)	NC(A)	log2FC	Padj	NC (G)	NC(A)
gyrA	DNA gyrase subunit A	-0.11	0.52	40895	44113	NA	NA	NA	NA	NA	NA	NA	NA
gyrB	DNA gyrase subunit B	-0.07	0.80	30889	32426	NA	NA	NA	NA	NA	NA	NA	NA
*TK0149**	Pyruvoyl-dependent arginine decarboxylase	-0.18	0.37	15065	17049	**-0.43**	**0.0052**	17049	23037	**-0.61**	**4.88E-6**	15065	23037
Rep74	Rolling circle replication initiator protein	**-0.36**	0.13	2149	2560	-0.05	0.9	2765	2954	**-0.41**	**0.046**	2149	2954
p24	Orphan DNA binding protein	0.16	0.48	196	220	**-0.67**	**5.77E-6**	195	312	**-0.51**	**0.00033**	196	312
Kan	Kanamycin resistance gene	0.28	0.49	79	65	**-1.01**	**0.0003**	65	132	**-0.73**	**0.0056**	79	132
Amp	Ampicillin resistance gene	0.18	0.44	232	204	-0.25	0.26	204	243	-0.06	0.79	232	243

NC (G) – normalised counts TKgyrAB strain; NC (Y) normalised counts TKY119F strain; NC (A) - normalised counts TKAg; mean value of four replicates is given. NA – non applicable

The values above the significance threshold are indicated in bold letters (Padj ≤ 0.05 and Log2FC ≥ |0.33|).

* confers agmatine prototrophy to *T. kodakarensis*

Supplementary table 5: Expression profile of topologically relevant genes

			TKgyrAB vs TKY119				**TKY119F vs TKAg**				**TKgyrAB vs TKAg**			
Locus name	Old locus name	Annotation	log2FC	Padj	NC (G)	NC(Y)	log2FC	Padj	NC (Y)	NC(A)	log2FC	Padj	NC (G)	NC(A)
TK_RS02320	*TK0470*	reverse gyrase	0.25	**0.007**	27209	22822	0.23	**0.023**	22822	19434	**0.49**	**1.6E-09**	27209	19434
TK_RS02325	*TK0471*	TrmBL2	-0.12	0.580	19845	21549	0.30	0.070	21549	17478	0.18	2.3E-01	19845	17478
TK_RS02760	*TK0560*	Alba	0.30	0.122	41197	33370	0.31	0.126	33370	26977	**0.61**	**7.9E-05**	41197	26977
TK_RS03845	*TK0778*	Mini-A	0.05	0.749	5065	4892	0.16	0.146	4892	4368	0.21	**1.9E-02**	5065	4368
TK_RS03950	*TK0798*	DNA topoisomerase VI subunit A	-0.004	0.980	7165	7185	0.15	**0.045**	7185	6459	0.15	**2.2E-02**	7165	6459
TK_RS03955	*TK0799*	DNA topoisomerase VI subunit B	0.02	0.887	8763	8617	0.09	0.467	8617	8086	0.12	2.3E-01	8763	8086
TK_RS05360	*TK1091*	DNA topoisomerase III	-0.04	0.772	6031	6212	0.19	0.061	6212	5445	0.15	1.0E-01	6031	5445
TK_RS07015	*TK1413*	histone A	-0.29	0.296	12472	15259	**0.41**	0.107	15259	11501	0.12	6.6E-01	12472	11501
TK_RS11530	*TK2289*	histone B	0.19	0.401	9589	8377	0.30	0.151	8377	6818	**0.49**	**2.2E-03**	9589	6818
TK_RS05005	*TK1017*	chromosome segregation protein SMC	-0.07	0.815	15166	15965	0.08	0.782	15965	15078	0.008	9.8E-01	15166	15078

NC (G) – normalised counts TKgyrAB strain; NC (Y) normalised counts TKY119F strain; NC (A) - normalised counts TKAg; mean value of four replicates is given.

The values above the significance threshold are indicated in bold letters (Padj $\leq$ 0.05 and Log2FC $\geq$ 1.25).

Supplementary table 6: Fragments per kilobase million counts for annotated topoisomerases in *T. kodakarensis*

			TKgyrAB (FPKM)				**TKY119F**				**TKAg**			
Locus name	Old locus name	Annotation	Rep1	Rep2	Rep3	Rep4	Rep1	Rep2	Rep3	Rep4	Rep1	Rep2	Rep3	Rep4
N.A.	*N.A.*	GyrA	**1.41**	**1.44**	**1.58**	**1.56**	1.43	1.51	1.68	1.79	*N.A.*	*N.A.*	*N.A.*	*N.A.*
N.A.	*N.A.*	GyrB	**0.57**	**0.59**	**0.65**	**0.72**	0.57	0.60	0.66	0.80	*N.A.*	*N.A.*	*N.A.*	*N.A.*
TK_RS02320	*TK0470*	reverse gyrase	**0.47**	**0.45**	**0.43**	**0.53**	0.41	0.37	0.39	0.39	0.38	0.32	0.33	0.36
TK_RS03845	*TK0778*	Mini-A	0.51	0.46	0.47	0.52	0.53	0.46	0.46	0.43	0.45	0.47	0.42	0.42
TK_RS03950	*TK0798*	DNA topoisomerase VI subunit A	0.55	0.53	0.55	0.55	0.55	0.54	0.54	0.55	0.54	0.53	0.51	0.48
TK_RS03955	*TK0799*	DNA topoisomerase VI subunit B	0.50	0.46	0.44	0.43	0.48	0.45	0.44	0.40	0.48	0.44	0.42	0.41
TK_RS05360	*TK1091*	DNA topoisomerase III	0.14	0.14	0.13	0.16	0.14	0.14	0.15	0.15	0.14	0.13	0.12	0.15

N.A. not applicable

REFERENCES

1. Taboada,B., Estrada,K., Ciria,R. and Merino,E. (2018) Operon-mapper: a web server for precise operon identification in bacterial and archaeal genomes. *Bioinformatics*, **34**, 4118–4120.

Article III

Reverse gyrase of *Thermococcus kodakarensis* has positive supercoiling activity *in vivo*

Paul Villain[1], Ryan Catchpole[1,2], Jacques Oberto[1], Patrick Forterre[1,3], Tamara Basta[1]

Affiliations:

1. Université Paris-Saclay, CEA, CNRS, Institute for Integrative Biology of the Cell (I2BC), 91198, Gif-sur-Yvette, France
2. Present address : Department of Biochemistry and Molecular Biology, University of Georgia, Athens, GA 30602, USA
3. Archaeal Virology Unit, Institut Pasteur, Paris, France

ABSTRACT

Hyperthermophiles, organisms living at temperatures above 80°C, have evolved a plethora of protective features helping them maintain genome stability in conditions in which DNA is subjected to high rates of DNA damage. Only one of these features, a topoisomerase reverse gyrase, is known to be uniquely found in all hyperthermophiles and many thermophiles suggesting an important role of this enzyme for adaptation to thermophily. Considerable effort has been made in the past 20 years to characterise the function of reverse gyrase both *in vivo* and *in vitro*. These studies resulted in many new and sometimes contradictory findings which, in a very resumed manner, point collectively to general function of this enzyme in DNA repair but also in controlling the global DNA supercoiling level in most hyperthermophilic archaea. In the present study, we characterized *in vivo* the impact of reverse gyrase deletion on DNA supercoiling in hyperthermophilic archaeon *Thermococcus kodakarensis*. We found that plasmid DNA was converted from positively supercoiled to relaxed topology in reverse gyrase deletion mutant thus demonstrating that reverse gyrase indeed functions as topoisomerase *in vivo*. The micrococcal digestion of chromatin did not reveal a detectable modification in the mutant *T. kodakarensis* cells suggesting that the lack of positive supercoiling activity has no effect on nucleosome assembly. We anticipate that the global scale studies using psoralen crosslink and MNase digestion coupled to next generation sequencing will allow us to determine without ambiguity how reverse gyrase affects the global supercoiling of the chromosome in hyperthermophiles.

INTRODUCTION

High temperatures are considerably challenging for cell physiology and particularly for genome stability. As the molecular motion increases proportionally to rise in temperature, molecules, membranes and DNA become unstable. Concretely, in *in vitro* conditions exposure of DNA to high temperature results in high rates of oxidation, alkylation, deamination, depurination and finally, strand breaks (Lindahl 1993; Marguet and Forterre 1994). To protect their genomes against these toxic DNA lesions, thermophiles and hyperthermophiles rely on specific adaptations. Among others, they harbour GC-rich rRNA, specific lipids, high proportions of thermostable amino acids in their proteins (typically I, V, Y, W R, E, L), chaperones, high salts and polyamines concentrations in their cytoplasms, specific ions transporters and a topoisomerase, the reverse gyrase (Grogan 1998; Zeldovich, Berezovsky, and Shakhnovich 2007; López-García et al. 2015).

This peculiar type IA topoisomerase was discovered in 1984 in *Sulfolobus acidocaldarius* and named after its unique positive supercoiling activity, that mirrored the one of DNA gyrase (Kikuchi and Asai 1984). Reverse gyrase is a fusion of a SF2-helicase and a topoisomerase domains, and consequently exhibits ATP-dependent and $MgCl_2$-dependent activity **(Figure 1)** (Confalonieri et al. 1993).

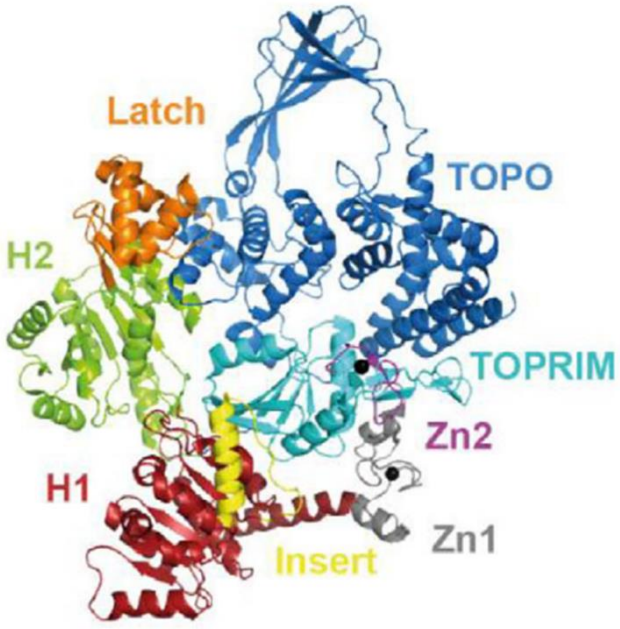


Figure 1: Structure of *Thermotoga maritima* reverse gyrase

Reverse gyrase is a fusion of the SF2 helicase domain and topoisomerase domain. The two domains cooperate thanks to the latch (orange). The helicase domain is composed of H1 (red), H2 (green), the zinc finger Zn1 (gray) and the insert (yellow). The topoisomerase domain is composed of the TOPO domain (dark blue), the TOPRIM domain (light blue) and another zinc finger Zn2 (purple) (adapted from McKie, Neuman, et Maxwell 2021).

It has been proposed that reverse gyrase creates both negative and positive supercoiling by opening DNA and then relaxing only the negative supercoiling leading to a net increase in positive supercoiling. The helicase and the topoisomerase domains cooperate thanks to the latch during the whole process (Lulchev and Klostermeier 2014). Remarkably, so far, reverse gyrase is the only protein shared by every hyperthermophile and most thermophiles, making it the hallmark of hyperthermophily (Forterre 2002). Its protective effect against high temperature was thought to come from the higher stability of positively supercoiled DNA, a topology that displays increased base stacking and is more resistant to double strand melting at high temperature.

The coating of DNA through chromatinization is another strategy to stabilize DNA at high temperature. Histones from archaea are classically organized in tetrameric nucleosomes (Mattiroli et al. 2017). But some archaea like *T. kodakarensis*, possess histones that can extend this basic structure by several dimers of histones (Maruyama et al. 2013). This results in assembly of histones into oligomers of various size that coat the DNA in a filamentous structure. *T. kodakarensis* was shown to exhibit almost complete histone coverage in its chromatin (Rojec et al. 2019) This high histone coverage over the genome suggests that histones are major actors of genome architecture and gene regulation in *T. kodakarensis*. Importantly, histones can titrate supercoiling both by constraining it or by releasing supercoils from opposite sense. Conversely, histones binding to DNA and histones oligomerization can be impacted by DNA supercoiling.

In most hyperthermophiles investigated so far plasmids were found in relaxed or slightly positively supercoiled state (López-García and Forterre 1999; Charbonnier and Forterre 1994). However, hyperthermophiles with negatively supercoiled DNA exist and they invariably encode DNA gyrase, indicating that when the two topoisomerases co-exist in a cell, the DNA gyrase dictates the overall DNA topology (Guipaud et al. 1997; Bouthier de la Tour et al. 1998; López-García et al. 2000). We recently demonstrated, by introducing an active DNA gyrase in the naïve hyperthermophilic archaeon *Thermococcus kodakarensis* (Villain et al. 2021), that the topology of plasmids can be converted from positive to negative supercoiling without adverse effect on growth of this archaeon. This demonstration further challenges the hypothesis that positively supercoiled DNA is an important

requirement for hyperthermophiles and raises questions about the physiological role of reverse gyrase.

Investigation of reverse gyrase function in hyperthermophilic archaea using mutational analyses produced ambiguous results. Deletion of the gene encoding reverse gyrase in euryarchaeon *T. kodakarensis* is viable and does not cause growth defect at 85°C, the optimal growth temperature. However, at 90°C, the growth rate is decreased by 2-fold and the strain does not grow at all at 93°C (Atomi, Matsumi, and Imanaka 2004). In closely related *Pyrococcus furiosus* the absence of reverse gyrase is lethal at 95°C, a suboptimal growth temperature for this archaeon which optimally grows at 100°C (Lipscomb et al. 2017). In crenarchaeon *Sulfolobus islandicus* each of the two genes encoding reverse gyrase can be deleted separately, but the deletion of TopR2 can only be obtained after 14-20 days of incubation that seems hard to reconcile with survival in natural growth conditions (Zhang et al. 2013; 2018).

Besides its positive supercoiling activity, several lines of evidence support direct involvement of reverse gyrase in DNA repair (reviewed in Garnier et al. 2021). Reverse gyrase of *Sulfolobus solfataricus* forms stable covalent complexes with UV-damaged DNA *in vitro* either using purified enzyme or with cell extracts (Napoli et al. 2004). At high protein/DNA ratio, the reverse gyrase of *Archaeoglobus fulgidus* prevents inappropriate aggregation of heat-denatured DNA regions and promotes correct annealing *in vitro* thus protecting DNA from heat-induced double strand breaks independently of any enzymatic activity (Kampmann and Stock 2004). Using global protein-interaction network approach the interaction was detected between reverse gyrase and the single strand DNA binding protein RPA of *Pyrococcus abyssi* (Pluchon et al., 2013). In *Sulfolobus solfataricus* reverse gyrase was found to form DNA-bridged complex with the single strand binding protein SSB (Napoli et al., 2005). The binding to SSB enhanced the topoisomerase activity of reverse gyrase suggesting its direct involvement in the DNA repair. In the following study the same group showed that reverse gyrase co-immunoprecipitates with translational polymerase Y (PolY) known for its essential role in DNA repair (Valenti et al. 2009). The reverse gyrase inhibits the activity of PolY *in vitro* and the positive supercoiling activity is required for inhibition suggesting that reverse gyrase may regulate the PolY activity when DNA damage occurs Other experiments examined the role of reverse gyrase when *Sulfolobus islandicus* cells were challenged with methyl methanesulfonate (MMS), a DNA alkylation agent which provokes DNA backbone breakage. The depletion of TopR1 in these conditions resulted in accelerated genomic DNA degradation accompanied by a higher rate of cell death (Han, Feng, and She 2017).

Since reverse gyrase was recognised as the only hallmark of hyperthermophily 20 years ago, a substantial amount of work was done to understand what makes this protein indispensable for life at

high temperature. In spite of this considerable effort the physiological role of reverse gyrase and in particular, its precise contribution to maintenance of genome integrity remains unclear. One of the main unresolved questions remains to understand what is the contribution of the reverse gyrase to the overall chromosomal DNA supercoiling. To address this question, we have characterized *in vivo* the impact of reverse gyrase deletion on cellular DNA topology. We first demonstrated that the reverse gyrase is indeed responsible for positive supercoiling of plasmids in hyperthermophilic archaeon *Thermococcus kodakarensis*. Using micrococcal nuclease (MNase) digestion of chromatin, we show that reverse gyrase deletion does not lead to aberrant profile suggesting that the lack of positive supercoiling activity in *T. kodakarensis* cells has no strong effect on nucleosome assembly.

MATERIAL AND METHODS

Strain construction

Transformation of *T. kodakarensis* was performed using standard protocol as previously described (Villain et al. 2021). The *T. kodakarensis* DAD (Δ*pyrF*, Δ*TK0149*) and DADΔRG (Δ*pyrF*, Δ*TK0149*, Δ*TK0470*) strains, are both auxotroph for uracil (Δ*pyrF*) and agmatine (Δ*TK0149*). DADΔRG carries, in addition, the deletion of the gene encoding reverse gyrase (Δ*TK0470*). The two strains were kindly provided by Hiroki Higashibata (Toyo University). A subsequent Δ*trpE*::*pyrF* replacement was made in these strains to allow tryptophan-based selection, resulting in TKDAD and TKΔRG strains. These two strains were transformed for two-dimensional gel electrophoresis with pTPTK3 replicating plasmid (Catchpole et al. 2018) resulting in TKDADp and TKΔRGp strains.

Two-dimensional agarose gel electrophoresis

To prepare samples for 2D agarose gel electrophoresis, Ravot medium was inoculated from overnight precultures of TKDADp and TKΔRGp at 1:100 dilution and incubated at 85°C without shaking until late exponential growth phase. Plasmids were gently extracted to preserve their topological integrity, as previously described (Villain et al. 2021).

Agarose gel electrophoresis was performed as previously described with minor modifications (Villain et al. 2021). Briefly, agarose gels were prepared by dissolving 0.8 % (w/v) of agarose in 1X TBE (89 mM Tris, 89 mM boric acid, 2 mM EDTA). Electrophoresis was performed at room temperature (usually 24°C) and minimum 2 µg of plasmid DNA was used for each sample. In the first dimension, no intercalating agent or 1.5 µg/ml of chloroquine was added in the gel and in the running buffer. Electrophoresis was run at 1.2 V/cm for 22 h. Gels were subsequently equilibrated for 1 h in 1X TBE buffer supplemented with 7.5 µg/ml of chloroquine and placed in the tray after a rotation of 90°. In the second dimension, an electric field of 2 V/cm was applied for 5 h. Gels were washed 3 x 30 min in water to remove chloroquine and then stained for 1h with 2.5 µg/ml of ethidium bromide. Gels were then rinsed with water and imaged with a Typhoon Imager using Cy3 channel. The main topoisomer from each plasmid preparation was determined by quantifying the intensity of individual topoisomer bands on gel images using Fiji software (Schindelin et al. 2012).

Nucleosome profiling by micrococcal nuclease digestion

Chromatin isolation and MNase digestion were carried out as previously described with minor modifications (Sanders et al. 2021). Briefly, strains TKDAD and TKΔRG were cultivated from fresh overnight precultures in 500 ml of Ravot medium until late exponential phase. Cultures were rapidly chilled in precooled beaker immersed in a water-ice bath to stop cellular metabolism and thus

minimize the risk of nucleosome disassembly. Cells were pelleted by centrifugation (4500 x g, 15 min, 4°C) and then washed with Ravot salt solution before storing at -80°C. Frozen pellets were resuspended in 1 ml of MNase buffer (50 mM Tris-HCl pH 8, 100 mM NaCl, and 1 mM $CaCl_2$) and then submitted to 5 cycles of nitrogen freezing plus grinding with mortar. Cell lysate was subsequently recovered by pipetting from the mortar and clarified by centrifugation for 5 min at 1,700 x g and 4°C. 500 µl of clarified lysates were carefully collected by pipetting and digested for 1 h at 37°C with 70 U of RNase A (Qiagen). Samples were separated into 5 x 100 µl fractions and then each fraction was digested 3 min at 37°C with either 0, 1, 10, 100 or 500 U of MNase (Thermo Scientific, ≥100 U/µl). DNA from digested fractions was immediately extracted with 300 µl of 10 mM Tris-HCl pH 8 and 400 µl of phenol/chloroform/isoamyl alcohol (25:24:1). Mixtures were vortexed for about 2 min and then centrifuged for 4 min at 14,000 x g, room temperature. 200 µl of the top aqueous phase were collected for each sample. DNA was precipitated by adding 200 µl of 1 M Tris-HCl pH 8 and 1 ml of pure ethanol, followed by overnight incubation at -80°C. DNA fragments were pelleted 30 min at 20,000 x g and 4°C. Resulting pellets were resuspended in 20 µl of TE (10 mM Tris-HCl pH 8, 1 mM EDTA) and resolved on a 4% agarose gel in 0.5X TAE buffer (20 mM Tris, 10 mM acetic acid, 0.5 mM EDTA) for 3 h at 4 V/cm.

RESULTS

In vivo measurement of reverse gyrase supercoiling activity in *Thermococcus kodakarensis*

The positive supercoiling of plasmids extracted from Thermococcales and other hyperthermophilic archaea is thought to come from reverse gyrase based on *in vitro* enzymatic assays either with purified protein or with cell extracts. However, the positive supercoiling activity of the reverse gyrase has never been tested *in vivo*. To this end, we used two-dimensional agarose gel electrophoresis to analyse the topology of plasmids extracted from strains TKDADp or TKΔRGp.

This allowed to determine the major topoisomers for reporter plasmid pTPTK3 isolated from TKDADp and TKΔRGp strains and calculate their superhelical density (σ) which (P. López-García and Forterre 1997). This value takes into account the size of the plasmid (4710 pb) and the effect of temperature shift from 85°C, the optimal growth temperature of *T. kodakarensis*, to 22°C the temperature of electrophoresis (~22°C) on plasmid topology (**Figure 2, A**). Superhelical density of σ = +0.0089 was determined for TKDADp strain, in good agreement with published values. In contrast, the superhelical density of pTPTK3 plasmid isolated from strain TKΔRGp drops to zero (**Figure 2, A**). Identical values of superhelical density were obtained when chloroquine was added in both dimensions (**Figure 2, B**).

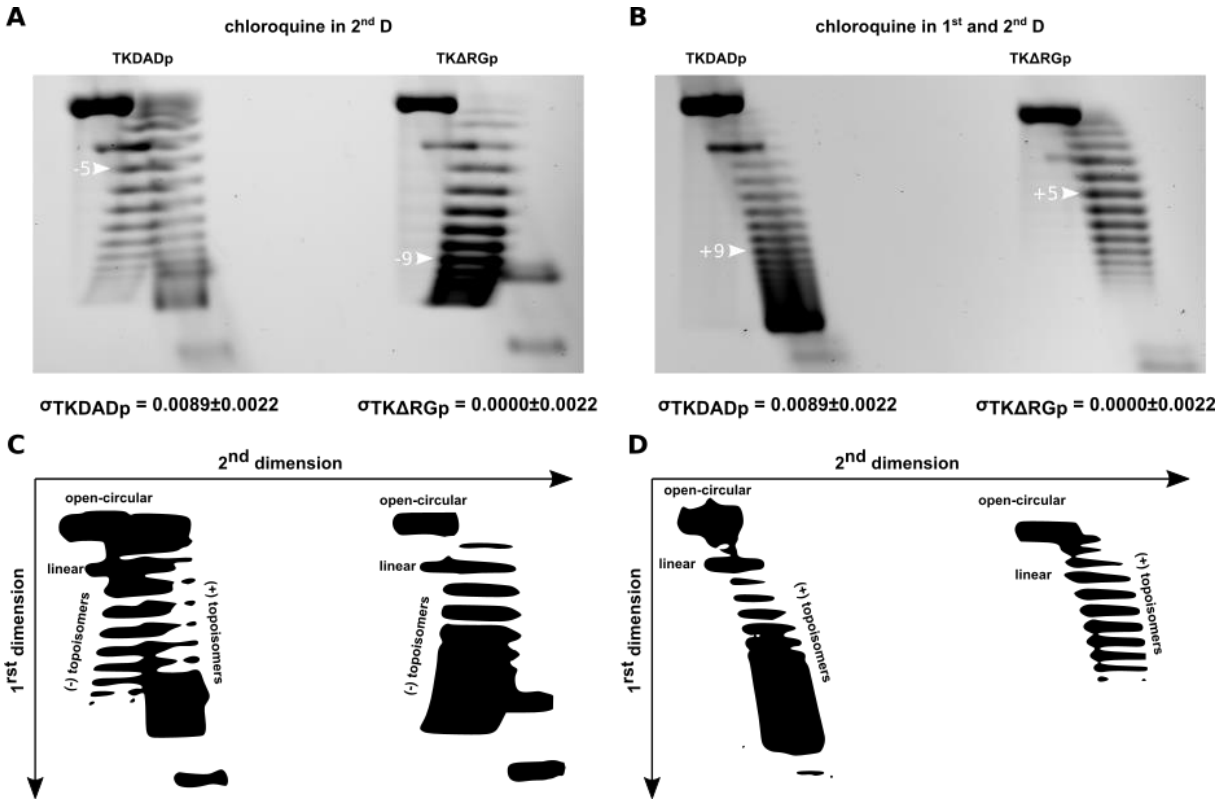


Figure 2: Plasmid DNA isolated from *Thermococcus kodakarensis* TKΔRGp is relaxed.

Reporter plasmid pTPTK3 was isolated from TKDADp or TKΔRGp strains and the topoisomers were resolved using 2D agarose gel electrophoresis. The corresponding cartoon image showing the distribution of different topological forms of pTPTK3 is depicted below each gel. Plasmid topoisomers were resolved by their degree of supercoiling in the first dimension, independently of their sense of supercoiling. In the second dimension, the addition of DNA intercalating drug chloroquine allows to discriminate between positive and negative supercoiling. Chloroquine introduces positive supercoils in plasmid DNA, thus increasing the electrophoretic mobility of positively supercoiled DNA or decreasing the migration of negatively supercoiled DNA. This results in a typical bell-shaped distribution, with negative, slow migrating topoisomers in the left arch and positive, fast migrating topoisomers in the right arch.

(A) and (C) DNA intercalating drug chloroquine was added only in the second dimension run to separate positive and negative topoisomers. Major topoisomers are indicated with a white arrow and these were used to calculate superhelical densities indicated below each gel. The error margin indicated (±0.0022) accounts for a potential mistake of ±1 during the determination of the main topoisomer.

(B) and (D) Chloroquine was added both in the first and the second dimension runs. The addition of the chloroquine in the first run allows to resolve topoisomers from TKΔRG strain by shifting them towards a positive supercoiling. Major topoisomers strain are indicated with a white arrow and used to calculate superhelical densities indicated below the gel. The error margin indicated (±0.0022) accounts for a potential mistake of ±1 during the determination of the main topoisomer.

Histone oligomerization is not affected in *T. kodakarensis* TKΔRG

We used micrococcal nuclease (MNase) digestion of chromatin to assess the oligomerization state of histones in nucleosomes. Using various concentration of MNase, we observed protected fragments of 60, 90, 120, 150 bp corresponding to oligomers of up to four dimers (**Figure 3)**. The most prominent band corresponds to 90 bp DNA fragments showing that trimer of histone dimers is the dominant oligomer in our experimental conditions (**Figure 3)**. Digestion patterns of TKDAD and TKΔRG strains were very similar, suggesting that reverse gyrase deletion does not affect histone oligomerization in *T. kodakarensis*.

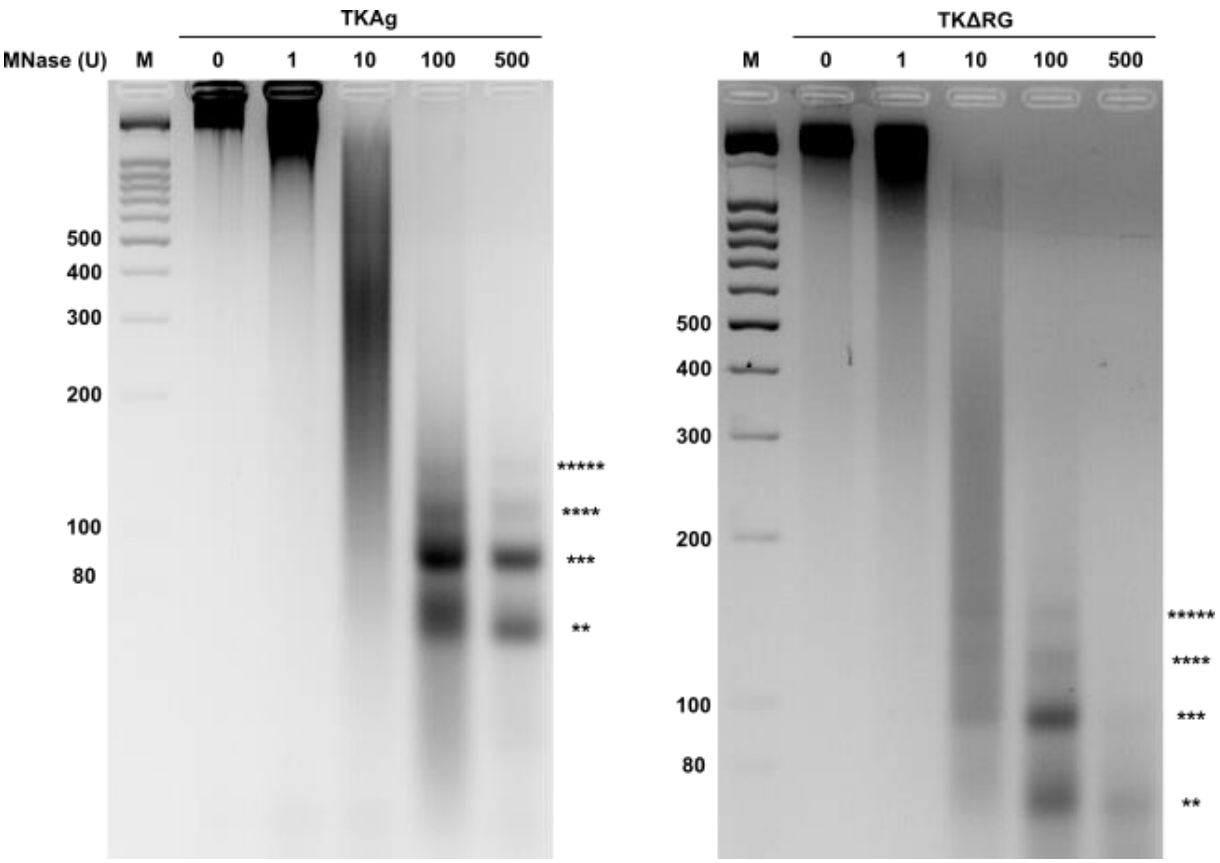


Figure 3: The lack of reverse gyrase has no major consequences for assembly of nucleosomes in *T. kodakarensis* TKΔRG

Digestion of chromatin isolated from TKAg (left panel) or TKΔRG (right panel) strain using increasing amounts of MNase. The resulting DNA fragments were separated on a 4% (w/v) agarose gel. The asterisks indicate the 60, 90, 120 and 150 bp bands corresponding to two, three, four or five histone dimers, respectively.

DISCUSSION

The precise role of reverse gyrase in hyperthermophilic organisms is still an open question. While positive supercoiling activity of reverse gyrase appears as an obvious candidate to protect genome against heat-induced damages, a substantial evidence suggests that this enzyme may also function as a DNA chaperone and in DNA repair. Besides this experimental evidence, the possibility for hyperthermophiles to harbour negatively supercoiled DNA argues strongly against positively supercoiled DNA as prerequisite for thermophilic lifestyle. (Guipaud et al. 1997; López-García et al. 2000). We recently reinforced this idea by showing that *T. kodakarensis* can accommodate a dominant gyrase activity, that changed its DNA topology from positive to negative supercoiling, without adverse effects upon growth (Villain et al. 2021). Noteworthy, the reverse gyrase was the sole endogenous topoisomerase that was upregulated in response to gyrase activity but this dysregulation was insufficient to restore the initial, positive, DNA topology in *T. kodakarensis*. These results suggested that the maintenance of a narrow range of supercoiling, compatible with DNA transactions, is not prerequisite for life in Thermococcus. This is in stark contrast with what is known for model bacteria where supercoiling levels are strictly regulated by joint action of DNA gyrase and Topo I and even slight perturbation of this system is lethal (Pruss, Manes, and Drlica 1982; Ferrándiz et al. 2016). In *T. kodakarensis*, reverse gyrase is an obvious candidate with the topoisomerase VI to handle the accumulation of negative supercoiling consecutive to DNA replication and transcription. In line with this hypothesis, it has been recently proposed that the TopR1reverse gyrase is important for supercoiling homeostasis in *Sulfolobus solfataricus* (Couturier et al. 2020). In addition, *topR2* was shown to be an essential gene in *S. solfataricus*, as expected for a gene involved in the removing of topological constrains generated by DNA transactions (Zhang et al. 2013; 2018). However, deletion of reverse gyrase is perfectly viable in *T. kodakarensis* at optimal growth temperature, indicating that replication and transcription are still possible without its positive supercoiling activity (Atomi, Matsumi, and Imanaka 2004).

In light on these data it remains unclear what is the function of reverse gyrase in controlling overall chromosomal supercoiling and whether this enzyme collaborates with topoisomerase VI in preventing stalling of DNA transaction processes because of accumulation of DNA supercoils. Thus, a comprehensive study *in vivo* of the exact role of reverse gyrase on DNA topology seems now essential to clarify this puzzling situation.

The work presented in this paper draft is a first attempt to initiate such a study. To the best of our knowledge, we performed the first *in vivo* measurement of reverse gyrase contribution to DNA topology. The results confirm that all the positive supercoiling detected by plasmid topoisomers

analysis can be attributed to reverse gyrase, excluding transcription-induced supercoiling. This results corroborate plasmid topology studies performed *in vitro* from cells extracts (Atomi, Matsumi, and Imanaka 2004) and confirm the high tolerance of *T. kodakarensis* to DNA topology variations as demonstrated in a previous study (Villain et al. 2021). Two-dimensional electrophoresis of small plasmid extracted directly from the organism studied is a reliable method to directly measure the supercoiling state of DNA. This will be a valuable tool to characterize further the importance of reverse gyrase on the DNA topology of hyperthermophiles.

Chromatin digestion by MNase shows the same typical protection pattern (60, 90, 120, 150 bp) corresponding to nucleosomes of up to five histones dimers in TKDAD and TKΔRG. This indicates that the loss of positive supercoiling consecutive to reverse gyrase deletion does not impact histones oligomerization. However, we did not observe longer fragments as reported in previous studies (Maruyama et al. 2011; 2013; Mattiroli et al. 2017; Sanders et al. 2021). This difference of histone oligomerization could be explained by a different growth stage of cells used to extract chromatin, or the different culture medium used in our study (Ravot medium instead of ASW-YT). The mapping of histone binding across the whole genome by MNase-seq should yield valuable information about the interplay between the reverse gyrase function and chromatinization in *T. kodakarensis*.

Together, this study and previous meticulous work accomplished to develop genetic engineering in archaea, paves the way for a more detailed investigation of the relation between histones and DNA supercoiling in hyperthermophilic archaea.

REFERENCES

Atomi, Haruyuki, Rie Matsumi, and Tadayuki Imanaka. 2004. 'Reverse Gyrase Is Not a Prerequisite for Hyperthermophilic Life'. *Journal of Bacteriology* 186 (14): 4829–33. https://doi.org/10.1128/JB.186.14.4829-4833.2004.

Bouthier de la Tour, C., C. Portemer, H. Kaltoum, and M. Duguet. 1998. 'Reverse Gyrase from the Hyperthermophilic Bacterium Thermotoga Maritima: Properties and Gene Structure'. *Journal of Bacteriology* 180 (2): 274–81. https://doi.org/10.1128/JB.180.2.274-281.1998.

Catchpole, Ryan, Aurore Gorlas, Jacques Oberto, and Patrick Forterre. 2018. 'A Series of New E. Coli-Thermococcus Shuttle Vectors Compatible with Previously Existing Vectors'. *Extremophiles: Life Under Extreme Conditions* 22 (4): 591–98. https://doi.org/10.1007/s00792-018-1019-6.

Charbonnier, F., and P. Forterre. 1994. 'Comparison of Plasmid DNA Topology among Mesophilic and Thermophilic Eubacteria and Archaebacteria'. *Journal of Bacteriology* 176 (5): 1251–59.

Confalonieri, F., C. Elie, M. Nadal, C. de La Tour, P. Forterre, and M. Duguet. 1993. 'Reverse Gyrase: A Helicase-like Domain and a Type I Topoisomerase in the Same Polypeptide'. *Proceedings of the National Academy of Sciences of the United States of America* 90 (10): 4753–57. https://doi.org/10.1073/pnas.90.10.4753.

Couturier, Mohea, Danièle Gadelle, Patrick Forterre, Marc Nadal, and Florence Garnier. 2020. 'The Reverse Gyrase TopR1 Is Responsible for the Homeostatic Control of DNA Supercoiling in the Hyperthermophilic Archaeon Sulfolobus Solfataricus'. *Molecular Microbiology* 113 (2): 356–68. https://doi.org/10.1111/mmi.14424.

Ferrándiz, María-José, Antonio J. Martín-Galiano, Cristina Arnanz, Isabel Camacho-Soguero, José-Manuel Tirado-Vélez, and Adela G. de la Campa. 2016. 'An Increase in Negative Supercoiling in Bacteria Reveals Topology-Reacting Gene Clusters and a Homeostatic Response Mediated by the DNA Topoisomerase I Gene'. *Nucleic Acids Research* 44 (15): 7292–7303. https://doi.org/10.1093/nar/gkw602.

Forterre, Patrick. 2002. 'A Hot Story from Comparative Genomics: Reverse Gyrase Is the Only Hyperthermophile-Specific Protein'. *Trends in Genetics: TIG* 18 (5): 236–37. https://doi.org/10.1016/s0168-9525(02)02650-1.

Garnier, Florence, Mohea Couturier, Hélène Débat, and Marc Nadal. 2021. 'Archaea: A Gold Mine for Topoisomerase Diversity'. *Frontiers in Microbiology* 12: 661411. https://doi.org/10.3389/fmicb.2021.661411.

Grogan, D. W. 1998. 'Hyperthermophiles and the Problem of DNA Instability'. *Molecular Microbiology* 28 (6): 1043–49. https://doi.org/10.1046/j.1365-2958.1998.00853.x.

Guipaud, O., E. Marguet, K. M. Noll, C. B. de la Tour, and P. Forterre. 1997. 'Both DNA Gyrase and Reverse Gyrase Are Present in the Hyperthermophilic Bacterium Thermotoga Maritima'. *Proceedings of the National Academy of Sciences of the United States of America* 94 (20): 10606–11.

Han, Wenyuan, Xu Feng, and Qunxin She. 2017. 'Reverse Gyrase Functions in Genome Integrity Maintenance by Protecting DNA Breaks In Vivo'. *International Journal of Molecular Sciences* 18 (7): E1340. https://doi.org/10.3390/ijms18071340.

Kampmann, Martin, and Daniela Stock. 2004. 'Reverse Gyrase Has Heat-Protective DNA Chaperone Activity Independent of Supercoiling'. *Nucleic Acids Research* 32 (12): 3537–45. https://doi.org/10.1093/nar/gkh683.

Kikuchi, A., and K. Asai. 1984. 'Reverse Gyrase--a Topoisomerase Which Introduces Positive Superhelical Turns into DNA'. *Nature* 309 (5970): 677–81.

Lindahl, T. 1993. 'Instability and Decay of the Primary Structure of DNA'. *Nature* 362 (6422): 709–15. https://doi.org/10.1038/362709a0.

Lipscomb, Gina L., Elin M. Hahn, Alexander T. Crowley, and Michael W. W. Adams. 2017. 'Reverse Gyrase Is Essential for Microbial Growth at 95 °C'. *Extremophiles: Life Under Extreme Conditions* 21 (3): 603–8. https://doi.org/10.1007/s00792-017-0929-z.

López-García, P., and P. Forterre. 1997. 'DNA Topology in Hyperthermophilic Archaea: Reference States and Their Variation with Growth Phase, Growth Temperature, and Temperature Stresses'. *Molecular Microbiology* 23 (6): 1267–79.

López-García, P., and P. Forterre. 1999. 'Control of DNA Topology during Thermal Stress in Hyperthermophilic Archaea: DNA Topoisomerase Levels, Activities and Induced Thermotolerance during Heat and Cold Shock in Sulfolobus'. *Molecular Microbiology* 33 (4): 766–77.

López-García, P., P. Forterre, J. van der Oost, and G. Erauso. 2000. 'Plasmid PGS5 from the Hyperthermophilic Archaeon Archaeoglobus Profundus Is Negatively Supercoiled'. *Journal of Bacteriology* 182 (17): 4998–5000. https://doi.org/10.1128/jb.182.17.4998-5000.2000.

López-García, Purificación, Yvan Zivanovic, Philippe Deschamps, and David Moreira. 2015. 'Bacterial Gene Import and Mesophilic Adaptation in Archaea'. *Nature Reviews. Microbiology* 13 (7): 447–56. https://doi.org/10.1038/nrmicro3485.

Lulchev, Pavel, and Dagmar Klostermeier. 2014. 'Reverse Gyrase—Recent Advances and Current Mechanistic Understanding of Positive DNA Supercoiling'. *Nucleic Acids Research* 42 (13): 8200–8213. https://doi.org/10.1093/nar/gku589.

Marguet, E., and P. Forterre. 1994. 'DNA Stability at Temperatures Typical for Hyperthermophiles'. *Nucleic Acids Research* 22 (9): 1681–86. https://doi.org/10.1093/nar/22.9.1681.

Maruyama, Hugo, Janet C Harwood, Karen M Moore, Konrad Paszkiewicz, Samuel C Durley, Hisanori Fukushima, Haruyuki Atomi, Kunio Takeyasu, and Nicholas A Kent. 2013. 'An Alternative Beads-on-a-String Chromatin Architecture in Thermococcus Kodakarensis'. *EMBO Reports* 14 (8): 711–17. https://doi.org/10.1038/embor.2013.94.

Maruyama, Hugo, Minsang Shin, Toshiyuki Oda, Rie Matsumi, Ryosuke L. Ohniwa, Takehiko Itoh, Katsuhiko Shirahige, et al. 2011. 'Histone and TK0471/TrmBL2 Form a Novel Heterogeneous Genome Architecture in the Hyperthermophilic Archaeon Thermococcus Kodakarensis'. *Molecular Biology of the Cell* 22 (3): 386–98. https://doi.org/10.1091/mbc.e10-08-0668.

Mattiroli, Francesca, Sudipta Bhattacharyya, Pamela N. Dyer, Alison E. White, Kathleen Sandman, Brett W. Burkhart, Kyle R. Byrne, et al. 2017. 'Structure of Histone-Based Chromatin in Archaea'. *Science (New York, N.Y.)* 357 (6351): 609–12. https://doi.org/10.1126/science.aaj1849.

McKie, Shannon J., Keir C. Neuman, and Anthony Maxwell. 2021. 'DNA Topoisomerases: Advances in Understanding of Cellular Roles and Multi-Protein Complexes via Structure-Function Analysis'. *BioEssays: News and Reviews in Molecular, Cellular and Developmental Biology*, January, e2000286. https://doi.org/10.1002/bies.202000286.

Napoli, Alessandra, Anna Valenti, Vincenzo Salerno, Marc Nadal, Florence Garnier, Mosè Rossi, and Maria Ciaramella. 2004. 'Reverse Gyrase Recruitment to DNA after UV Light Irradiation in Sulfolobus Solfataricus'. *The Journal of Biological Chemistry* 279 (32): 33192–98. https://doi.org/10.1074/jbc.M402619200.

Pruss, G. J., S. H. Manes, and K. Drlica. 1982. 'Escherichia Coli DNA Topoisomerase I Mutants: Increased Supercoiling Is Corrected by Mutations near Gyrase Genes'. *Cell* 31 (1): 35–42. https://doi.org/10.1016/0092-8674(82)90402-0.

Rojec, Maria, Antoine Hocher, Kathryn M. Stevens, Matthias Merkenschlager, and Tobias Warnecke. 2019. 'Chromatinization of Escherichia Coli with Archaeal Histones'. *ELife* 8. https://doi.org/10.7554/eLife.49038.

Sanders, Travis J., Fahad Ullah, Alexandra M. Gehring, Brett W. Burkhart, Robert L. Vickerman, Sudili Fernando, Andrew F. Gardner, Asa Ben-Hur, and Thomas J. Santangelo. 2021. 'Extended Archaeal Histone-Based Chromatin Structure Regulates Global Gene Expression in Thermococcus Kodakarensis'. *Frontiers in Microbiology* 12: 681150. https://doi.org/10.3389/fmicb.2021.681150.

Schindelin, Johannes, Ignacio Arganda-Carreras, Erwin Frise, Verena Kaynig, Mark Longair, Tobias Pietzsch, Stephan Preibisch, et al. 2012. 'Fiji: An Open-Source Platform for Biological-Image Analysis'. *Nature Methods* 9 (7): 676–82. https://doi.org/10.1038/nmeth.2019.

Valenti, Anna, Giuseppe Perugino, Takehiko Nohmi, Mosè Rossi, and Maria Ciaramella. 2009. 'Inhibition of Translesion DNA Polymerase by Archaeal Reverse Gyrase'. *Nucleic Acids Research* 37 (13): 4287–95. https://doi.org/10.1093/nar/gkp386.

Villain, Paul, Violette Da Cunha, Etienne Villain, Patrick Forterre, Jacques Oberto, Ryan Catchpole, and Tamara Basta. 2021. 'The Hyperthermophilic Archaeon Thermococcus Kodakarensis Is Resistant to Pervasive Negative Supercoiling Activity of DNA Gyrase', 16.

Zeldovich, Konstantin B., Igor N. Berezovsky, and Eugene I. Shakhnovich. 2007. 'Protein and DNA Sequence Determinants of Thermophilic Adaptation'. *PLoS Computational Biology* 3 (1): e5. https://doi.org/10.1371/journal.pcbi.0030005.

Zhang, Changyi, Alex P. R. Phillips, Rebecca L. Wipfler, Gary J. Olsen, and Rachel J. Whitaker. 2018. 'The Essential Genome of the Crenarchaeal Model Sulfolobus Islandicus'. *Nature Communications* 9 (1): 4908. https://doi.org/10.1038/s41467-018-07379-4.

Zhang, Changyi, Bin Tian, Suming Li, Xiang Ao, Kevin Dalgaard, Serkan Gökce, Yunxiang Liang, and Qunxin She. 2013. 'Genetic Manipulation in Sulfolobus Islandicus and Functional Analysis of DNA Repair Genes'. *Biochemical Society Transactions* 41 (1): 405–10. https://doi.org/10.1042/BST20120285.

Article IV

Genome wide mapping of DNA supercoiling in hyperthermophilic archaeon *Thermococcus kodakarensis*

Paul Villain[1], Ryan Catchpole[1,2], Jacques Oberto[1], Patrick Forterre[1,3] & Tamara Basta[1]

Affiliations:

1. Université Paris-Saclay, CEA, CNRS, Institute for Integrative Biology of the Cell (I2BC), 91198, Gif-sur-Yvette, France

2. Present address : Department of Biochemistry and Molecular Biology, University of Georgia, Athens, GA 30602, USA

3. Archaeal Virology Unit, Institut Pasteur, Paris,

ABSTRACT

In all cells the DNA is present in topologically constrained, supercoiled, state. The vast majority of organisms have negatively supercoiled DNA and they need the strand-opening potential of negative DNA supercoiling to allow transcription and other DNA-dependent processes. The only exception to this rule are some hyperthermophilic archaea which harbour a topoisomerase reverse gyrase that positively supercoiles DNA. The importance of negative DNA supercoiling for controlling gene expression in model bacteria and eukaryotes is now well established notably thanks to the development of techniques for genome-wide mapping of supercoiling. In this work we set up one of these techniques, psoralen photocrosslink, to investigate the distribution of supercoiling in hyperthermophilic archaeon *Thermococcus kodakarensis* which naturally has slightly positively supercoiled DNA. We also applied this technique to *T. kodakarensis* in which DNA was converted to negatively supercoiled DNA by expressing bacterial DNA gyrase and to *T. kodakarensis* with relaxed DNA due to deletion of reverse gyrase. We found that psoralen can be efficiently crosslinked to *T. kodakarensis* DNA *in vivo* under optimal growth conditions of this organism. Under low hit conditions (ratio crosslinked/non-crosslinked DNA = 0.2) we did not see significant difference in amount of crosslinked DNA between different strains suggesting that the global amount of negative supercoiling remained similar and that differences, if they exist, may occur locally on their genomes. The MNase digestion profile of chromatin did not show significant difference between the different strains suggesting that introduced topological perturbations do not affect histone oligomerisation.

INTRODUCTION

The helical nature of DNA while providing stability to the molecule, imposes strong topological constrains every time that the two strands must be separated (Watson and Crick 1953). This means that DNA transactions like replication, transcription, recombination and repair complexify DNA topology, resulting in supercoils, knots, catenanes and non-B DNA transitions (Liu and Wang 1987; Postow et al. 2001; Bates and Maxwell 2005). The accumulation of these structures influence in return DNA transactions, interfering with normal cellular processes. To maintain topological constrains at a physiological level, cells rely on a wide set of actors dedicated to DNA topology that cooperate to release constrains or contain them in specific regions of chromosomes. The most common and important type of DNA topological structure is supercoiling, which is directly regulated by topoisomerases (Schoeffler and Berger 2008; Wang 2002; Bush, Evans-Roberts, and Maxwell 2015; Seol and Neuman 2016; McKie, Neuman, and Maxwell 2021).

DNA supercoiling has been extensively studied in model organisms by studying plasmid topology using sucrose gradient or agarose gel electrophoresis. These methods are based on the increased mobility of supercoiled plasmids (as compared to relaxed plasmids) due to their compactness. While easy to set up, these methods are limited by the fact that plasmids, due to their small size, do not faithfully represent much bigger and complex chromosomes. Genetic approaches, relying on the site-specific recombination of resolvase γδ or bacteriophage λ integrase (Bliska and Cozzarelli 1987; Cibot, n.d.), was used in the past to estimate supercoiling distribution across the chromosome. However, these methods are restricted to genetically tractable model bacteria and imply genetic engineering with potential side effects. Other methods based on psoralen derivatives overcome these limitations by allowing measurement of relative supercoiling directly on chromosomes without need for genetic engineering (Corless and Gilbert 2017). Psoralen is a little planar aromatic molecule that intercalates between DNA base pairs and establishes crosslink under UV light. Underwound DNA is about two fold more likely to incorporate psoralen, making it a reliable tool for mapping negative DNA supercoiling (Sinden, Bat, and Kramer 1999; Bermúdez et al. 2010). In the last decade, psoralen has been used to investigate supercoiling at the genome scale in various model bacteria and eukaryotes (Bermúdez et al. 2010; Teves and Henikoff 2014; Kouzine et al. 2013; Naughton et al. 2013; Lal et al. 2016; Achar et al. 2020) but not in Archaea. Consequently, nothing is known about the supercoiling state of archaeal chromosomes.

Negative supercoiling of DNA has a crucial role in the regulation of DNA transactions by its potential to decrease base stacking interactions, thus favouring double helix opening. Positive supercoiling of DNA on the other hand stabilizes double helix in a closed conformation decreasing its melting potential. In

Bacteria, supercoiling is used as a quick and massive transcriptional regulator by its ability to change simultaneously the expression of large number of genes in response to environmental changes (Rui and Tse-Dinh 2003; Peter et al. 2004; Dorman and Dorman 2016, Villain et al. 2021). This regulation is mainly based on the DNA gyrase activity that introduces negative supercoiling in the genome proportionally to the [ATP]/[ADP] ratio in the cell (Westerhoff et al. 1988). Consequently, bacterial genomes are in average underwound but the distribution of the negative supercoiling is not homogenous and varies with respect to origin of replication and the growth phase (Worcel and Burgi 1972; Lal et al. 2016). In *Escherichia coli*, the only bacterium for which the supercoiling has been mapped genome wide, a decreasing gradient of negative supercoiling is measured from the terminus to the origin of replication in stationary phase while no gradient was observed when the bacterium grew exponentially. This specific pattern was shown to correlate with transcription intensity and gyrase binding sites (Sobetzko, Travers, and Muskhelishvili 2012; Sutormin et al. 2018), reflecting the importance of this topoisomerase for efficient transcription and global control of supercoiling. Eukaryotes lack DNA gyrase but they also have underwound genome (Sinden, Carlson, and Pettijohn 1980). This topology is achieved by wrapping DNA around nucleosomes which cover the majority of eukaryotic genomes. The gene expression control from such constrained loci is generated by a complex cross-talk between transcription, histones binding and topoisomerase relaxation (Baranello et al. 2012). Importantly, in Eukaryotes, free negative supercoils are concentrated near transcriptional start sites (TSS) and transcriptional termination sites (TTS) of active genes, while free positive supercoils can be found in middle of actively transcribed genes (Achar et al. 2020). Such an organization promotes non-B DNA structures at genes boundaries that insulate the gene from the rest of the genome, thus increasing RNA polymerase recycling from TSS to TTS and keeping gene bodies overwound to avoid nucleosome repositioning.

The studies of global DNA topology in Archaea are still in their infancy but initial studies of *in vivo* plasmid topology indicate that, as it is often the case, these organisms harbour idiosyncratic features not found in bacteria or eukaryotes. For example, positively supercoiled plasmid DNA is so far only found in thermophilic archaeal lineages (Charbonnier et al. 1992; Charbonnier and Forterre 1994; López-García and Forterre 1997; López-García et al. 2000). This specificity is thought to protect genome integrity at high temperatures, but it probably also affects global DNA geometry and gene expression regulation. Beyond the special case of thermophiles, Archaea possess small and dense genomes with the majority of genes structured in operons and their transcription and translation are coupled. These bacterial-like properties are counterbalanced by a transcription machinery homologous to the eukaryotic one and the presence of histones in most archaeal linages (Sanders, Marshall, and Santangelo 2019). It is probable that successful integration in course of evolution of such mosaic

combination of molecular features resulted in emergence of original mechanisms for control of DNA topology in Archaea.

In this study, we used trimethylpsoralen (TMP) crosslink coupled to next-generation sequencing (NGS) to establish the first map of DNA supercoiling in an archaeon. We focused on *Thermococcus kodakarensis*, a genetically tractable hyperthermophilic archaeon with a relatively simple topological kit composed of two histones (HtkA and HtkB), two NAPs (TrmBL2 and Alba), one SMC and three topoisomerases (TopoIII, TopoVI and reverse gyrase). We used four previously constructed strains: a reference TKAg strain, a control strain TKY119F expressing inactive gyrase, a gyrase expressing strain TKgyrAB and a reverse gyrase deletion mutant TKΔRG (Villain et al. 2021). By comparing relative TMP crosslink enrichment on the genome of these strains, we plan to (i) map the distribution of negative supercoiling in *T. kodakarensis*, giving insight on interplay between local DNA topology, gene transcription and higher-order folding of chromosome (ii) establish directly the importance of reverse gyrase for global chromosome supercoiling; (iii) describe in more detail the impact of gyrase acquisition on the chromosome topology of a naïve organism.

MATERIAL AND METHODS

Thermococcus kodakarensis strains and culture

T. kodakarensis strains TKAg, TKY119F, TKgyrAB and TKΔRG were cultured in anaerobic conditions in ravot medium at 85°C as described previously (Villain et al. 2021). The TKgyrAB strain encodes the gyrase of *Thermotoga maritima* on the pTNAg plasmid, the strain TKY119F encodes a catalytic mutant of the gyrase and the TKAg strain encodes the empty pTNAg plasmid. The TKΔRG is a reverse gyrase deletion mutant was generously supplied by Hiroki Higashibata (Tokyo university).

In vivo plasmid DNA crosslinking with psoralen

Trimethylpsoralen (Sigma-Aldrich, ref. T6137) was diluted to saturation (0.9 mg/ml) in absolute ethanol. 50 µg/ml of psoralen were injected in sealed bottles containing 20 ml of late exponential phase cultures of *T. kodakarensis* strains TKY119F and TKgyrAB grown at 85°C. To avoid cold-shock, injections were performed directly in the oven. After 30 min of incubation at 85°C with the TMP, cultures were poured in the lid of a 10 cm glass Petri dish maintained at 85°C using a water-bath. Cells were immediately irradiated with ~9.6 $kJ.m^{-2}.min^{-1}$ of 365-nm light using a 45 W lamp (Vilber Lourmat model VL-315.BL). After irradiation, cells were immediately chilled on ice before to be pelleted by 15 min of centrifugation at 5,000 x g and at 4°C. Control samples not treated with TMP or treated with TMP but not irradiated with UV were treated in the same way. Plasmid DNA was extracted using the NucleoSpin Plasmid kit (Macherey Nagel) following the manufacturer's protocol. Then, 600 ng of plasmid per sample was digested 40 min at 37°C with FastDigest BglII restriction enzyme (ThermoFisher). Digested DNA was purified using Nucleospin Gel and PCR kit (Macherey Nagel) following the manufacturer's protocol. Final elution was done with 30 µl NE buffer (5 mM Tris/HCl, pH 8.5). Half of each plasmid sample was heated 5 min at 95°C in a thermocycler to denature DNA and then immediately chilled on ice. Samples were analysed by electrophoresis on 1% agarose gel in 0.5X TAE buffer (20 mM Tris, 10 mM acetic acid, 0.5 mM EDTA). Gels were stained in a 3X Gelred (Biotium) bath and imaged with a Typhoon Imager (Amersham) using Cy3 channel.

In vivo chromosomal DNA crosslinking with psoralen

The protocol for psoralen crosslinking of *T. kodakarensis* DNA is based on the protocol published by Kouzine and collaborators (Drolet 2018) with modifications. Trimethylpsoralen (Sigma-Aldrich, ref. T6137) was diluted up to saturation (0.9 mg/ml) in absolute ethanol. For initial screening of

appropriate crosslinking conditions various amounts of the TMP solution or equivalent volume of ethanol (negative controls), were injected in sealed bottles containing 20 ml of late exponential phase *T. kodakarensis* strains grown at 85°C. The concentration of 1 µg/ml was chosen for preparing samples for sequencing. To avoid cold-shock, injection was performed directly in the oven. After 10 min of incubation at 85°C with the TMP, cultures were poured in the lid of a 10 cm glass Petri dish maintained at 85°C using a water-bath. Cells were immediately irradiated with ~9.6 $kJ.m^{-2}.min^{-1}$ of 365-nm light using a 45 W lamp (Vilber Lourmat model VL-315.BL). For the sequenced samples were irradiated by this way 3 min (28.8 kJ/m^2). The anaerobic state of cultures during irradiation was controlled through the resazurin indicator contained in the Ravot medium. No yellow to pink transition, synonym of medium oxidation, was observed for irradiation times inferior to 3 min. After irradiation, cells were immediately chilled on ice before harvesting at 4°C by 15 min of centrifugation at 5,000 x g.

Total DNA extraction and fragmentation

Pellets were resuspended in 250 µl of TEN (40 mM Tris-HCl pH 7.5, 1 mM EDTA, 150 mM NaCl) and then lysed by the addition of 250 µl TENST (40 mM Tris-HCl pH 7.5, 1 mM EDTA, 150 mM NaCl, 1.6% N-lauryl sarcosine, 0.12% Triton X-100). Cell lysates were digested with 0.5 mg of RNAse A (Qiagen) for 1 h at 37°C and then incubated overnight at 55°C with 0.5 mg of proteinase K (Thermofisher Scientific). DNA was extracted by classical phenol/chloroform method, with three rounds of phenol/chloroform/isoamyl alcohol extraction (Sigma-Aldrich, 25:24:1) and one chloroform/isoamyl alcohol extraction (VWR, Ready-Red). Aqueous phase was carefully recovered and DNA precipitated by addition of 0.8 volume of isopropanol. After 1 h of incubation at -80°C, DNA was pelleted at 20,000 x g, 4°C for 30 min. Dried pellets were resuspended in 85 µl of 5 mM Tris/HCl, pH 8.5. DNA was fragmented to an average size of 200 bp by 200 cycles of 180 s of sonication with 10% duty and 175 peak, using a S220 focused-ultrasonicator (Covaris).

Separation of TMP crosslinked and non-crosslinked DNA fragments

The sonicated DNA was separated on a 0.65% agarose gel in 0.5X TAE buffer. DNA fragments of 100-300 bp were cut out of agarose gel and the DNA was extracted with the NucleoSpin Gel and PCR Clean-up kit (Macherey-Nagel) and eluted with 15 µl of ultrapure water pH 8.5. DNA was not exposed to UV light during the gel excision. After addition of 4 µl of 0.1 M phosphate buffer (57.7 mM Na_2HPO_4 and 42.3 mM NaH_2PO_4, pH 7), DNA samples were heat-denatured for 5 min at 100°C, then 20 µl of a mix of DMSO and glyoxal (Alfa Aesar) was immediately added (respective final concentrations of 40%

and 4%) followed by 90 min of incubation at 55°C. Crosslinked and non-crosslinked DNA fragments were separated by electrophoresis at 2 V/cm for 14 h, with buffer recirculation between electrodes, on 3% agarose gels. Gels were 20 cm long and made of 1.2:1.8 mixture of UltraPure agarose and NuSieve 3:1 agarose in 200 ml of 10 mM phosphate buffer pH 7. TMP-DNA crosslinks were reversed by incubating gels 3 h at 65°C in denaturing solution (0.5M NaOH, 1.5M NaCl) followed by 2 h of incubation with agitation in neutralizing solution (1.5 M NaCl, 500 mM Tris–HCl pH 8, 1 mM EDTA). Then gels were quickly rinsed in MilliQ water before to be equilibrated 3 times 1 h in 1X TAE and stained overnight with SYBR Gold diluted in 1X TAE. Gels were then imaged with a Typhoon Imager (Amersham) using Cy2 channel. Pieces of gels containing crosslinked and non-crosslinked DNA fragments were excised with a clean scalpel and gel-extracted using a Gel and PCR Clean-up Midi kit with NTC buffer (Macherey-Nagel). DNA was eluted with 5 mM Tris/HCl and stored at -80°C before sequencing.

Nucleosome profiling by micrococcal nuclease digestion

Chromatin isolation and MNase digestion were carried out as previously described (Sanders et al. 2021) with minor modifications. Briefly, strains TKAg, TKY119F, TKgyrAB and TKΔRG were cultivated from fresh overnight precultures in 500 ml of Ravot medium until late exponential phase. Cultures were rapidly chilled in precooled beaker immersed in a water-ice bath to stop cellular metabolism and thus minimize the risk of nucleosome disassembly. Cells were pelleted by centrifugation (4500 x g, 15 min, 4°C) and then washed with Ravot salt solution before storing at -80°C. Frozen pellets were resuspended in 1 ml of MNase buffer (50 mM Tris-HCl pH 8, 100 mM NaCl, and 1 mM $CaCl_2$) and then submitted to 5 cycles of nitrogen freezing plus grinding with mortar. Cell lysate was subsequently recovered by pipetting from the mortar and clarified by centrifugation for 5 min at 1,700 x g and 4°C. 500 µl of clarified lysates were carefully collected by pipetting and digested for 1 h at 37°C with 70 U of RNase A (Qiagen). Samples were separated into 5 x 100 µl fractions and then each fraction was digested 3 min at 37°C with either 0, 50, 100, 200 or 400 U of MNase (Thermo Scientific, ≥100 U/µl). DNA from digested fractions was immediately extracted with 300 µl of 10 mM Tris-HCl pH 8 and 400 µl of phenol/chloroform/isoamyl alcohol (25:24:1). Mixtures were vortexed for about 2 min and then centrifuged for 4 min at 14,000 x g, room temperature. 200 µl of the top aqueous phase were collected for each sample. DNA was precipitated by adding 200 µl of 1 M Tris-HCl pH 8 and 1 ml of pure ethanol, followed by overnight incubation at -80°C. DNA fragments were pelleted 30 min at 20,000 x g and 4°C. Resulting pellets were resuspended in 20 µl of TE (10 mM Tris-HCl pH 8, 1 mM EDTA) and resolved on a 4% agarose gel in 0.5X TAE buffer (20 mM Tris, 10 mM acetic acid, 0.5 mM EDTA) for 2 h at 4 V/cm.

RESULTS

TMP can be crosslinked *in vivo* to DNA of *T. kodakarensis*

To get a snapshot of DNA topology across the whole genome, TMP assay must be done *in vivo*, preferentially at optimal culture conditions. This is particularly challenging in *T. kodakarensis* considering that it grows optimally at 85°C and in absence of oxygen. To test the feasibility of psoralen crosslink in this archaeon we incubated *T. kodakarensis* cells for 30 min with saturating concentration of TMP at 85°C in the reductive culture medium. Cells were then irradiated with UV to achieve crosslink and plasmid DNA was isolated. After digestion with restriction enzymes, plasmids were heat-denatured and analysed by electrophoresis. Only the samples treated with psoralen and irradiated with UV, exhibited the expected BglII digestion pattern (**Figure 1**). The presence of such pattern is only possible if plasmid DNA is covalently crosslinked to TMP and resists heat-denaturation. The data thus demonstrate that *in vivo* psoralen-DNA crosslink can be achieved under the conditions which preserve the native DNA topology of *T. kodakarensis*.

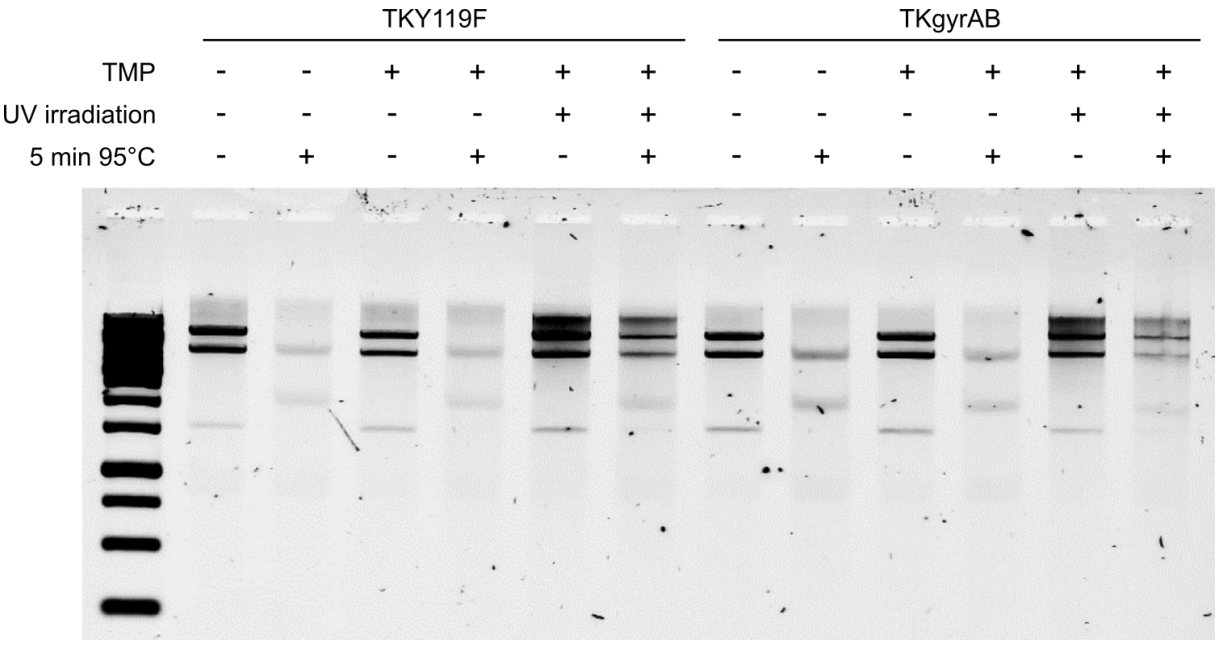


Figure 1: Plasmid DNA and TMP can be UV-crosslinked in cells of in *T. kodakarensis*

T. kodakarensis strains TKY119F and TKgyrAB were treated with TMP, irradiated with UV and their plasmids extracted. After digestion with BglII, the crosslinking of TMP into DNA was assessed by thermal treatment and gel agarose electrophoresis. Crosslinked DNA fragments were resistant to thermal denaturation and exhibit a 6787/4295/1520 bp digestion pattern similar to the unheated samples. Conversely, control samples without TMP treatment or without UV irradiation do not exhibit the expected pattern because the DNA is heat-denatured.

Set up of TMP crosslinking conditions for chromosomal DNA supercoiling mapping

In order to be used as a topological probe, the number of TMP-DNA crosslinks must follow linear correlation with the UV irradiation dose and the quantity of TMP used to treat the cells. The range of UV irradiation producing linear incorporation of DNA crosslinks was determined by *in vivo* TMP-crosslinking experiments. Cells were treated with 5 µg/ml of TMP and irradiated with increasing dose of UV. The ratio of crosslinked versus non-crosslinked DNA was then assessed by electrophoresis after glyoxal denaturation. The proportion of crosslinked DNA increased in a linear manner over the entire range of UV irradiation (from 10 s to 240 s at ~9.6 $kJ.m^{-2}.min^{-1}$) (**Figure 2A and B**).

Next, we investigated the dose-response of the TMP concentration on the quantity of DNA-crosslinks. Cells were treated with increasing amounts of TMP and irradiated for 180 s at ~9.6 $kJ.m^{-2}.min^{-1}$. Crosslinked and non-crosslinked DNA was separated by agarose gel electrophoresis after glyoxal denaturation. The proportion of crosslinked DNA increased in a linear manner up to 5 µg/ml of TMP. Beyond this concentration, the proportion of crosslinked DNA stagnated between 22% and 24% (**Figure 2C and D**).

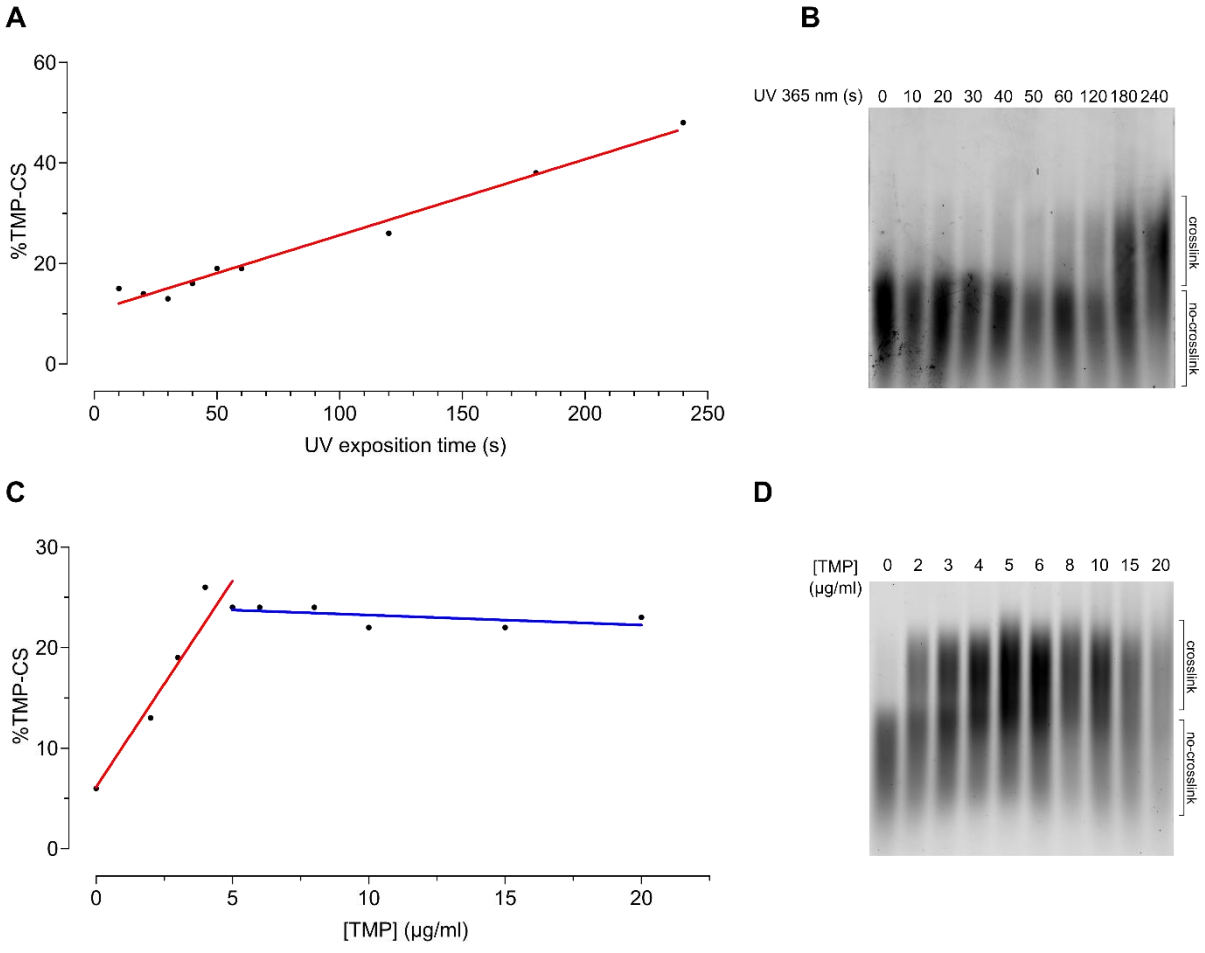


Figure 2: ***I****n vivo* crosslinking of TMP in DNA of *T. kodakarensis*.

TMP UV-crosslinking is an efficient probe for *in vivo* DNA supercoiling investigation in *T. kodakarensis* (A) TMP crosslink incorporation is linear over 240 s of 365 nm UV exposition at 9.6 $kJ.m^{-2}.min^{-1}$ (red line) (B) TMP crosslink incorporation is linear up to a 5 µg/ml concentration in the culture medium (red line), a saturation plateau is reached beyond this limit (blue line).

From these data we chose unsaturating TMP crosslink conditions (1 µg/ml of TMP and 3 min of UV irradiation) to prepare samples for DNA supercoiling mapping (**Figure 3**).

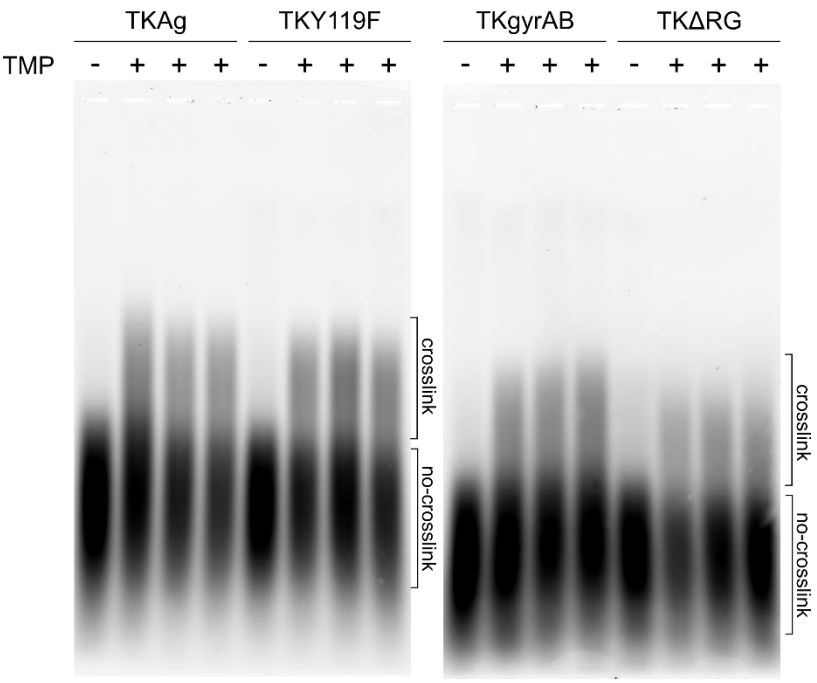


Figure 3: TMP sequencing samples

Crosslinked and non-crosslinked DNA fragments before gel excision and sequencing. Strain names, control lanes with no TMP treatment (-) and biological replicates (+) are indicated on the top of the gel.

For the sequenced samples, the ratio crosslinked/non crosslinked DNA was quantified by image analysis before gel extraction. Values were systematically inferior to 20% for each strain, meaning that TMP photobinding produced less than one crosslink per kb. In these low hit conditions, TMP-crosslinking is suitable to investigate DNA supercoiling with minor interference from DNA-binding proteins.

MNase digestion of chromatin

TMP intercalation into DNA can be affected by histones and nucleoid associated proteins (NAPs). To limit this bias, we used low hit conditions for TMP crosslinking. In this way, the incorporation of TMP into DNA is mainly due to helical tension (Bermúdez et al. 2010). Still, the native chromatin organization can impact the diffusion of supercoils and the activity of topoisomerases. To account for this phenomenon we performed MNase-seq experiment. We first extracted chromatin from the four *T. kodakarensis* strains and digested it using MNase.

The protected DNA fragments of 60/90/120/150 bp corresponding to two, three, four or five histone dimers, respectively were visible on agarose gels and their stoichiometry remained similar for all strains (**Figure 4**). This suggested that oligomerization of histones into polymeric nucleosomes was not strongly affected. Based on the digestion profile, we have chosen the samples digested with 200U of MNase as suitable for sequencing of MNase protected fragments. The corresponding four DNA bands were gel extracted and sent for DNA sequencing.

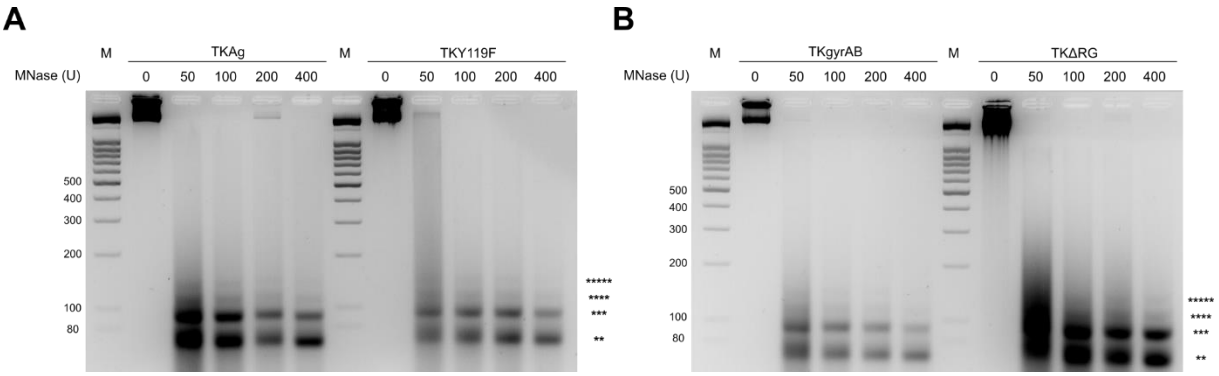


Figure 4: MNase-seq samples

Chromatin digestion with increasing amounts of MNase. The resulting DNA fragments were separated on a 4% (w/v) agarose gel. The asterisks indicate the 60, 90, 120 and 150 bp bands corresponding to two, three, four or five histone dimers, respectively. Samples were prepared from three biological replicates. The gels presented here are representative of the replicates.

DISCUSSION

The use of psoralen derivatives for genome wide mapping of supercoiling had produced major results over the last decades. Coupled to microarrays or next generation sequencing, this technique revealed the distribution of supercoiling within large topological domains down to precise genes or promoters in some bacteria and eukaryotes. We applied this approach to *T. kodakarensis*, to perform the first mapping of DNA supercoiling in an archaeon and in a hyperthermophile.

We found that the maximal TMP concentration producing linear increase of DNA crosslinks was 2 to 4 fold lower than those previously described for bacteria and eukaryotes. This may reflect an existence of a particular DNA topology in *T. kodakarensis* and/or efficient crossing of psoralen through the cellular envelope. In the same experimental conditions, the crosslink/non-crosslink ratio is similar for the four analysed strains. This suggests that the global amount of negative supercoiling is, in average, the same in all strains and that the differences, if any, may lie in the local distribution of negative supercoiling across the genome. The capacity of histones to assemble in polymers also seems not to be affected in strains that express gyrase (excess negative supercoiling) or that lack reverse gyrase (no positive supercoiling activity). The mapping of negative supercoiling is required to interpret this result since, again, the effects on chromatin distribution and structure may be local rather than global.

In conclusion, the sequencing results of TMP treated samples, coupled to the MNase-seq results, when available should give first insight of the general organization of DNA supercoiling in *T. kodakarensis*. The use of TKΔRG and TKgyrAB strains in this study should give us precious information about the importance of reverse gyrase in DNA topology regulation as well as the consequences of the introduction of pervasive negative supercoiling activity to chromosomal DNA topology of *T. kodakarensis*.

References

Achar, Yathish Jagadheesh, Mohamood Adhil, Ramveer Choudhary, Nick Gilbert, and Marco Foiani. 2020. 'Negative Supercoil at Gene Boundaries Modulates Gene Topology'. *Nature* 577 (7792): 701–5. https://doi.org/10.1038/s41586-020-1934-4.

Baranello, Laura, David Levens, Ashutosh Gupta, and Fedor Kouzine. 2012. 'The Importance of Being Supercoiled: How DNA Mechanics Regulate Dynamic Processes'. *Biochimica et Biophysica Acta (BBA) - Gene Regulatory Mechanisms*, Chromatin in time and space, 1819 (7): 632–38. https://doi.org/10.1016/j.bbagrm.2011.12.007.

Bates, Andrew D., and Anthony Maxwell. 2005. 'DNA Topology'. 2005. https://books.google.fr/books/about/DNA_Topology.html?id=WGBAGyzvQOUC&redir_esc=y.

Bermúdez, Ignacio, José García-Martínez, José E. Pérez-Ortín, and Joaquim Roca. 2010. 'A Method for Genome-Wide Analysis of DNA Helical Tension by Means of Psoralen-DNA Photobinding'. *Nucleic Acids Research* 38 (19): e182. https://doi.org/10.1093/nar/gkq687.

Bliska, J. B., and N. R. Cozzarelli. 1987. 'Use of Site-Specific Recombination as a Probe of DNA Structure and Metabolism in Vivo'. *Journal of Molecular Biology* 194 (2): 205–18. https://doi.org/10.1016/0022-2836(87)90369-x.

Bush, Natassja G., Katherine Evans-Roberts, and Anthony Maxwell. 2015. 'DNA Topoisomerases'. *EcoSal Plus* 6 (2). https://doi.org/10.1128/ecosalplus.ESP-0010-2014.

Charbonnier, F., G. Erauso, T. Barbeyron, D. Prieur, and P. Forterre. 1992. 'Evidence That a Plasmid from a Hyperthermophilic Archaebacterium Is Relaxed at Physiological Temperatures'. *Journal of Bacteriology* 174 (19): 6103–8.

Charbonnier, F., and P. Forterre. 1994. 'Comparison of Plasmid DNA Topology among Mesophilic and Thermophilic Eubacteria and Archaebacteria'. *Journal of Bacteriology* 176 (5): 1251–59.

Cibot, Camille. n.d. 'Etude du surenroulement diffusible de l'ADN chromosomique chez la bactérie Escherichia Coli', 238.

Corless, Samuel, and Nick Gilbert. 2017. 'Investigating DNA Supercoiling in Eukaryotic Genomes'. *Briefings in Functional Genomics* 16 (6): 379–89. https://doi.org/10.1093/bfgp/elx007.

Dorman, Charles J., and Matthew J. Dorman. 2016. 'DNA Supercoiling Is a Fundamental Regulatory Principle in the Control of Bacterial Gene Expression'. *Biophysical Reviews* 8 (3): 209–20. https://doi.org/10.1007/s12551-016-0205-y.

Drolet, Marc, ed. 2018. *DNA Topoisomerases: Methods and Protocols*. Methods in Molecular Biology. Humana Press. //www.springer.com/us/book/9781493974580.

Kouzine, Fedor, Ashutosh Gupta, Laura Baranello, Damian Wojtowicz, Khadija Ben-Aissa, Juhong Liu, Teresa M Przytycka, and David Levens. 2013. 'Transcription-Dependent Dynamic Supercoiling Is a Short-Range Genomic Force'. *Nature Structural & Molecular Biology* 20 (3): 396–403. https://doi.org/10.1038/nsmb.2517.

Lal, Avantika, Amlanjyoti Dhar, Andrei Trostel, Fedor Kouzine, Aswin S. N. Seshasayee, and Sankar Adhya. 2016. 'Genome Scale Patterns of Supercoiling in a Bacterial Chromosome'. *Nature Communications* 7 (March). https://doi.org/10.1038/ncomms11055.

Liu, L. F., and J. C. Wang. 1987. 'Supercoiling of the DNA Template during Transcription'. *Proceedings of the National Academy of Sciences of the United States of America* 84 (20): 7024–27.

López-García, P., and P. Forterre. 1997. 'DNA Topology in Hyperthermophilic Archaea: Reference States and Their Variation with Growth Phase, Growth Temperature, and Temperature Stresses'. *Molecular Microbiology* 23 (6): 1267–79.

López-García, P., P. Forterre, J. van der Oost, and G. Erauso. 2000. 'Plasmid PGS5 from the Hyperthermophilic Archaeon Archaeoglobus Profundus Is Negatively Supercoiled'. *Journal of Bacteriology* 182 (17): 4998–5000. https://doi.org/10.1128/jb.182.17.4998-5000.2000.

McKie, Shannon J., Keir C. Neuman, and Anthony Maxwell. 2021. 'DNA Topoisomerases: Advances in Understanding of Cellular Roles and Multi-Protein Complexes via Structure-Function Analysis'. *BioEssays: News and Reviews in Molecular, Cellular and Developmental Biology* 43 (4): e2000286. https://doi.org/10.1002/bies.202000286.

Naughton, Catherine, Nicolaos Avlonitis, Samuel Corless, James G. Prendergast, Ioulia K. Mati, Paul P. Eijk, Scott L. Cockroft, Mark Bradley, Bauke Ylstra, and Nick Gilbert. 2013. 'Transcription Forms and Remodels Supercoiling Domains Unfolding Large-Scale Chromatin Structures'. *Nature Structural & Molecular Biology* 20 (3): 387–95. https://doi.org/10.1038/nsmb.2509.

Peter, Brian J., Javier Arsuaga, Adam M. Breier, Arkady B. Khodursky, Patrick O. Brown, and Nicholas R. Cozzarelli. 2004. 'Genomic Transcriptional Response to Loss of Chromosomal Supercoiling in Escherichia Coli'. *Genome Biology* 5 (11): R87. https://doi.org/10.1186/gb-2004-5-11-r87.

Postow, L., N. J. Crisona, B. J. Peter, C. D. Hardy, and N. R. Cozzarelli. 2001. 'Topological Challenges to DNA Replication: Conformations at the Fork'. *Proceedings of the National Academy of Sciences of the United States of America* 98 (15): 8219–26. https://doi.org/10.1073/pnas.111006998.

Rui, Shan, and Yuk-Ching Tse-Dinh. 2003. 'Topoisomerase Function during Bacterial Responses to Environmental Challenge'. *Frontiers in Bioscience: A Journal and Virtual Library* 8 (January): d256-263. https://doi.org/10.2741/984.

Sanders, Travis J., Craig J. Marshall, and Thomas J. Santangelo. 2019. 'The Role of Archaeal Chromatin in Transcription'. *Journal of Molecular Biology*, May. https://doi.org/10.1016/j.jmb.2019.05.006.

Sanders, Travis J., Fahad Ullah, Alexandra M. Gehring, Brett W. Burkhart, Robert L. Vickerman, Sudili Fernando, Andrew F. Gardner, Asa Ben-Hur, and Thomas J. Santangelo. 2021. 'Extended Archaeal Histone-Based Chromatin Structure Regulates Global Gene Expression in Thermococcus Kodakarensis'. *Frontiers in Microbiology* 12: 681150. https://doi.org/10.3389/fmicb.2021.681150.

Schoeffler, Allyn J., and James M. Berger. 2008. 'DNA Topoisomerases: Harnessing and Constraining Energy to Govern Chromosome Topology'. *Quarterly Reviews of Biophysics* 41 (1): 41–101. https://doi.org/10.1017/S003358350800468X.

Seol, Yeonee, and Keir C. Neuman. 2016. 'The Dynamic Interplay Between DNA Topoisomerases and DNA Topology'. *Biophysical Reviews* 8 (3): 221–31. https://doi.org/10.1007/s12551-016-0206-x.

Sinden, R. R., O. Bat, and P. R. Kramer. 1999. 'Psoralen Cross-Linking as Probe of Torsional Tension and Topological Domain Size in Vivo'. *Methods (San Diego, Calif.)* 17 (2): 112–24. https://doi.org/10.1006/meth.1998.0723.

Sinden, R. R., J. O. Carlson, and D. E. Pettijohn. 1980. 'Torsional Tension in the DNA Double Helix Measured with Trimethylpsoralen in Living E. Coli Cells: Analogous Measurements in Insect and Human Cells'. *Cell* 21 (3): 773–83.

Sobetzko, Patrick, Andrew Travers, and Georgi Muskhelishvili. 2012. 'Gene Order and Chromosome Dynamics Coordinate Spatiotemporal Gene Expression during the Bacterial Growth Cycle'. *Proceedings of the National Academy of Sciences of the United States of America* 109 (2): E42-50. https://doi.org/10.1073/pnas.1108229109.

Sutormin, Dmitry, Natalia Rubanova, Maria Logacheva, Dmitry Ghilarov, and Konstantin Severinov. 2018. 'Single-Nucleotide-Resolution Mapping of DNA Gyrase Cleavage Sites across the Escherichia Coli Genome'. *Nucleic Acids Research*, December. https://doi.org/10.1093/nar/gky1222.

Teves, Sheila S., and Steven Henikoff. 2014. 'Transcription-Generated Torsional Stress Destabilizes Nucleosomes'. *Nature Structural & Molecular Biology* 21 (1): 88–94. https://doi.org/10.1038/nsmb.2723.

Villain, Paul, Violette Da Cunha, Etienne Villain, Patrick Forterre, Jacques Oberto, Ryan Catchpole, and Tamara Basta. 2021. 'The Hyperthermophilic Archaeon Thermococcus Kodakarensis Is Resistant to Pervasive Negative Supercoiling Activity of DNA Gyrase', 16.

Wang, James C. 2002. ‘Cellular Roles of DNA Topoisomerases: A Molecular Perspective’. *Nature Reviews. Molecular Cell Biology* 3 (6): 430–40. https://doi.org/10.1038/nrm831.
Watson, J. D., and F. H. C. Crick. 1953. ‘Molecular Structure of Nucleic Acids: A Structure for Deoxyribose Nucleic Acid’. *Nature* 171 (4356): 737–38. https://doi.org/10.1038/171737a0.
Westerhoff, H. V., M. H. O’Dea, A. Maxwell, and M. Gellert. 1988. ‘DNA Supercoiling by DNA Gyrase. A Static Head Analysis’. *Cell Biophysics* 12 (June): 157–81. https://doi.org/10.1007/BF02918357.
Worcel, A., and E. Burgi. 1972. ‘On the Structure of the Folded Chromosome of Escherichia Coli’. *Journal of Molecular Biology* 71 (2): 127–47. https://doi.org/10.1016/0022-2836(72)90342-7.

Conclusions and Perspectives

I. Evolutive history of gyrase in Archaea

Gyrase is ubiquitous in Bacteria but mainly restricted to Cluster II Euryarchaeota in Archaea. This distribution was thought to reflect a single horizontal gene transfer from Bacteria to Archaea leading to the emergence of mesophilic lifestyle in Archaea (Raymann et al. 2014; López-García et al. 2015). But the discovery of many new archaeal lineages over the last decade, complexified this view by revealing a wider distribution of gyrase (Adam et al. 2017). To clarify the evolution of gyrase in archaea, we reconstructed gyrase phylogeny in prokaryotes, using an up to date taxonomic sampling including the newly discovered archaeal lineages.

We conclude that the gyrase is present in three major archaeal phyla, Euryarchaea, DPANN and Asgard but never found in TACK archaea. The topology of gyrase tree recapitulates faithfully the consensus phylogeny of bacteria and archaea including the recovery of the two major bacterial clades Terrabacteria and Gracilicutes. Such topology indicates predominantly vertical evolution of gyrase. Our results are compatible with a single transfer of gyrase genes from bacteria to the ancestor of Cluster II Euryarchaea and subsequent dissemination via only few secondary gene transfers early through diversification of Archaea. From there we conclude that the establishment of gyrase in a naïve organism is not common in agreement with profound impact of this enzyme on essential cellular DNA transactions. We also propose that the gyrase phylogeny can help establish the temporal order of emergence of major bacterial and archaeal lineages, thus helping understand the most distant evolution of organisms. Finally, the short branches separating archaeal and bacterial clades are compatible with the emergence of the gyrase post LUCA in the lineage leading to bacteria.

To complete the general view of gyrase distribution in Archaea, it would be worth identifying the key determinants preventing or promoting its successful establishment in archaeal lineages. Assuming that most organisms rely on tightly regulated DNA topology to keep their genomes functional, chromatin architecture proteins like histones, NAPs and SMC as well as topoisomerases are probably critical in the establishment of gyrase in a given lineage. For instance, histones by constraining the excess of negative supercoiling generated by gyrase, can prevent or at least attenuate its deleterious consequences. From this respect, the absence of histones in Crenarchaeota and Diaforarchaea (part of TACK superphylum) could explain why gyrase is never found in these clades (Hocher et al. 2021).

It could be also envisaged that gyrase activity interferes with the activity of the topoisomerase VI. This type II topoisomerase, ubiquitous in Archaea, is probably in charge of positive supercoiling relaxation and chromosomes decatenation in Archaea, in a reminiscent way to gyrase and topoisomerase IV, respectively, in Bacteria. Thus, the introduction of gyrase into a naïve archaeon could give raise to conflicts with endogenous topoisomerase VI. Such conflicts could be attenuated by adaptation of gyrase and/or topoisomerase VI to new repartition of the workload and if so, this should be detectable at the sequence level. The phylogeny of the topoisomerase VI could reveal such phenomenon. In addition, comparing the activity and the structure of topoisomerase VI from gyrase-encoding archaea with those from gyrase-less archaea could give additional clues. As a particular evolutionary case, the Thermoplasmata, who are devoid of topoisomerase VI but encode gyrase, are of special interest. It is probable that their gyrase in addition to control of chromosomal supercoiling also ensures genome decatenation thus constituting a good model to understand adaptations to this specific situation.

II. Consequences of gyrase acquisition by a naïve archaeon

The development of reliable genetic tools in *T. kodakarensis* gave us the opportunity to describe the consequences of gyrase acquisition by a naïve organism. Besides, this organism is a good model mimicking the ancestral state of Euryarchaeota such that we could recapitulate as faithfully as possible the ancestral event of gyrase emergence in archaeal domain. We investigated the DNA topology, the growth and the transcriptome of the *T. kodakarensis* gyrase expressing strain.

We successfully introduced an active gyrase in a hyperthermophilic archaeon, and we thus conclude that such acquisition is possible even when the recipient is a modern organism. We also confirmed that gyrase can access the chromatinized genome of *T. kodakarensis* and that its supercoiling activity deregulates massively the expression of genes (about 10% of all *T. kodakarensis* annotated genes), suggesting that most of these genes are supercoiling sensitive. Despite a change of DNA topology from positive (σ = +0.0077) to negative (σ = -0.0327), we did not observe marked growth defect for the gyrase expressing strain. Since the level of supercoiling is tightly regulated in Bacteria and can undergo only little changes (about 1.1 – 1.4 fold) without toxicity, the 5-fold increase in negative supercoiling in *T. kodakarensis* is impressive. Our data also show that the endogenous topoisomerases are not used by *T. kodakarensis* to re-establish the native level of supercoiling thus suggesting an existence of an alternative mechanism for control of DNA supercoiling.

We propose that this mechanism could rely on high histone coverage observed in this organism with dense chromatin constraining most of the artificially introduced negative supercoiling, protecting the genome from excessive heat-induced melting and preserving DNA transactions.

The main conclusion that we draw from the introduction of gyrase in *T. kodakarensis* is that, in marked contrast to what is known for bacteria, DNA supercoiling does not need to be strictly maintained within a tiny range for cell survival. This unusual resistance of *T. kodakarensis* to DNA supercoiling variations could be explained by a buffering effect of histones. As *T. kodakarensis* histones deletion mutants or mutants of histones oligomerization are viable, it would be interesting to use them as recipient of gyrase to test this hypothesis. In parallel, the introduction of gyrase in *Sulfolobus* crenarchaeon, a thermophile without histones, would inform us how this organism would deal with alteration of DNA topology in absence of chromatin. Another pertinent model organism would be *Archaeoglobus* that are the only hyperthermophilic archaea that naturally encode gyrase. Unfortunately, these peculiar archaea, while cultivable, are not genetically tractable for the moment.

An attractive hypothesis is that gyrase was determinant for transition from thermophilic to mesophilic lifestyle. We conducted preliminary experiments that showed no obvious growth advantage of the gyrase expressing strain at sub-optimal temperature except perhaps at early stages of growth (the lag phase). However, over a long-term culturing experiment it would be interesting to test if synergic mutations would emerge to promote growth at low temperatures. This would be a major finding enhancing our understanding of adaptation to mesophilic lifestyle in course of evolution.

III. Importance of DNA supercoiling for adaptation to hyperthermophilic lifestyle

Thermophilic lifestyle requires a complex set of adaptations to ensure proper cell functioning. Among these adaptations, reverse gyrase is the only one common to all hyperthermophiles and most thermophiles but absent from mesophiles. By its positive supercoiling activity and its involvement in DNA repair, reverse gyrase contributes to

genome stability but ambiguous results about the relative importance of these two activities coupled to a lack of *in vivo* data, make its precise role elusive. Consequently, we initiated an *in vivo* characterization of reverse gyrase activities to clarify their importance in hyperthermophilic lifestyle. Moreover, histones and NAPs are surely major actors contributing to genome stability at high temperature, making them obvious partners of reverse gyrase. To better understand the relation between these topological actors, we included chromatin structure analysis in this study.

Using plasmid topology analysis, we demonstrated that the positive supercoiling observed in *T. kodakarensis* can be attributed to the reverse gyrase. Since deletion of reverse gyrase is not associated with growth defect at optimal growth temperature, we also conclude that the topoisomerase activity of reverse gyrase is not necessary for relaxing negative supercoil build-up during DNA transactions in *T. kodakarensis*. Moreover, the preliminary characterisation of chromatin structure showed that the loss of positive supercoiling has no dramatic impact on the oligomerization and DNA binding of histones.

The *in vivo* demonstration that reverse gyrase accounts for all the positive supercoiling in *T. kodakarensis* cells, suggests that the positive supercoiling deficit cannot be compensated by DNA transactions. Together with the viability of reverse gyrase deletion mutant, it reinforces the idea that positive supercoiling has a minor role in genome stability at high temperature. However, it could be envisaged that the reverse gyrase supercoiling activity is crucial outside of the "comfortable" culture conditions we used. To explore this hypothesis it would be interesting to investigate DNA topology in both the reverse gyrase deletion mutant and the wild type strain in stressful conditions, like changing temperatures or in presence of DNA-damaging agents. We initiated, in collaboration with A. Gardner's laboratory (New England Biolabs) a mirror study to map DNA damages genome-wide in reverse-gyrase deletion mutant. This should tell us if reverse gyrase protects the genome from DNA damage and if so what type of damage.

IV. Mapping DNA supercoiling along chromosomes by psoralen crosslinking

The characterization of DNA topology *in vivo* is hindered by the complexity, the diversity and the imbrication of the involved actors and mechanisms. In addition, DNA topology is highly dynamic and its supercoiling component can diffuse over long distances along the chromosome. For these reasons, but also because it is technically difficult to faithfully describe chromosomal DNA topology, analyses are often restricted to plasmids. By their circular shape and their reasonable sizes, plasmids appear as convenient tool to study DNA topology. However, their simplicity is also their weakness as they can not account for the complexity of chromosomes. This limitation was progressively lifted by the development of psoralen-based *in vivo* supercoiling measurement methods. There are different ways to proceed, but the basic principle of these methods is to exploit the propensity of psoralen derivatives to intercalate preferentially in negatively supercoiled DNA and to create inter-strand crosslinks under UV irradiation. Crosslinked and non-crosslinked DNA are then separated and sequenced to estimate the relative rate of negative supercoiling over the entire genome. While difficult to implement, these methods are actually the only currently existing allowing a direct measurement of negative supercoiling at the entire genome scale. I adapted psoralen crosslinking protocol originally developed for eukaryotes to *T. kodakarensis* and I used it to map DNA supercoiling on its chromosome. The sequencing data are now available and need to be analysed. The results of this analysis, when available, will be precious to understand the impact of the gyrase introduction on DNA supercoiling and gene expression as well as to decipher the exact contribution of reverse gyrase to global DNA supercoiling in *T. kodakarensis*. More generally, the mapping of DNA supercoiling in *T. kodakarensis*, opens the way for a general characterization of DNA topology regulation in this archaeon.

I used a psoralen derivative called trimethyl-psoralen (TMP) to map DNA supercoiling over the genome of *T. kodakarensis*. The method used is adapted from the protocol published by Fedor Kouzine and collaborators (Kouzine et al. 2013, 2018). It relies on the crosslinking of supercoiled DNA by 365 nm photobound TMP and its separation from non-crosslinked DNA by DNA denaturation and agarose gel electrophoresis. Populations of DNA are then sequenced and the crosslinked enrichment reflecting DNA supercoiling mapped on the genome.

The adaptation of this method to *T. kodakarensis* impose strong constrains due to the growth conditions of this archaeon. As DNA topology is extremely sensitive to environmental conditions, special attention was put to keep *T. kodakarensis* under anaerobiosis and at a temperature of 85°C, during the TMP treatment and the UV irradiation. Under these harsh conditions we were concerned about the ability of TMP to be stable, crosslinked by UV and specific of DNA supercoiling (absence of heat-induced non-specific crosslinks). Thus, preliminaries experiments of TMP *in vivo* crosslinking were conducted to set up conditions adapted to these constrains. They demonstrated that (i) TMP can enter *T. kodakarensis* cells, (ii) TMP is stable at 85°C during the ~15 min of the experiment, (iii) the temperature is not causing off target TMP intercalation into DNA, (iv) TMP crosslinking can be achieved by 365 nm irradiation in the growth conditions of *T. kodakarensis*.

In order to be used as a probe of DNA supercoiling, the relations between TMP crosslinking into DNA and the quantity of TMP used, as well as the dose of UVs used for the crosslinking, must be linear. To check for these conditions, I titrated the incorporation of TMP at constant irradiation conditions and defined this linear range for *T. kodakarensis*. It showed a linear incorporation of TMP in *T. kodakarensis* up to 5 µg/ml, a relatively low value compared to the thresholds described in *E. coli* or model eukaryotes. In addition, I demonstrated that TMP incorporation into *T. kodakarensis* DNA is linear over 4 min of UV irradiation at ~9.6 kJ/m^2/min.

The intercalation of TMP into DNA can be affected by the chromatin structure. This phenomenon is only revealed at saturating TMP concentrations. To limit this bias, experiments were carried using low, non-saturating concentrations of TMP to produce limited amount of crosslink (ratio crosslink / not crosslink < 1/3). A MNase-seq experiment was performed in parallel to confirm that potential variations of TMP incorporation can not be attributed to nucleosome redistribution over the genome.

The successful extraction of TMP crosslinked DNA that, *a priori*, fulfils all the conditions for the investigation of DNA supercoiling by TMP incorporation confirms the viability of this approach for *T. kodakarensis*.

The development of the psoralen photo-binding method in *T. kodakarensis* is a major step on the characterization of DNA topology organization and regulation in Archaea and hyperthermophiles. Coupled to the genetic editing developed for this organism, it opens the way for a global and integrative characterization of its DNA topology. Concretely, the use of this method on *T. kodakarensis* should give us the first view of DNA supercoiling organization in an archaeon, allowing comparison with the known situation of *E. coli* or eukaryotes. Moreover, this method would help to complete the data obtained on the introduction of a pervasive gyrase activity in *T. kodakarensis*, by giving a detailed view of the supercoiling perturbations at the chromosome scale. In the same idea, the comparison of DNA supercoiling density over the chromosomes of wild type and reverse gyrase deletion mutant strains, could be decisive to understand the importance of this peculiar enzyme on the stability of genomes of hyperthermophiles.

References

Adam, Panagiotis S., Guillaume Borrel, Céline Brochier-Armanet, and Simonetta Gribaldo. 2017. 'The Growing Tree of Archaea: New Perspectives on Their Diversity, Evolution and Ecology'. *The ISME Journal* 11 (11): 2407–25. https://doi.org/10.1038/ismej.2017.122.

Albers, S. V., J. L. Van de Vossenberg, A. J. Driessen, and W. N. Konings. 2001. 'Bioenergetics and Solute Uptake under Extreme Conditions'. *Extremophiles: Life Under Extreme Conditions* 5 (5): 285–94. https://doi.org/10.1007/s007920100214.

Albers, Sonja-Verena, Patrick Forterre, David Prangishvili, and Christa Schleper. 2013. 'The Legacy of Carl Woese and Wolfram Zillig: From Phylogeny to Landmark Discoveries'. *Nature Reviews. Microbiology* 11 (10): 713–19. https://doi.org/10.1038/nrmicro3124.

Ammar, Ron, Dax Torti, Kyle Tsui, Marinella Gebbia, Tanja Durbic, Gary D. Bader, Guri Giaever, and Corey Nislow. 2012. 'Chromatin Is an Ancient Innovation Conserved between Archaea and Eukarya'. *ELife* 1 (December): e00078. https://doi.org/10.7554/eLife.00078.

Arents, G., R. W. Burlingame, B. C. Wang, W. E. Love, and E. N. Moudrianakis. 1991. 'The Nucleosomal Core Histone Octamer at 3.1 A Resolution: A Tripartite Protein Assembly and a Left-Handed Superhelix'. *Proceedings of the National Academy of Sciences of the United States of America* 88 (22): 10148–52. https://doi.org/10.1073/pnas.88.22.10148.

Ashley, Rachel E., Andrew Dittmore, Sylvia A. McPherson, Charles L. Turnbough, Keir C. Neuman, and Neil Osheroff. 2017. 'Activities of Gyrase and Topoisomerase IV on Positively Supercoiled DNA'. *Nucleic Acids Research* 45 (16): 9611–24. https://doi.org/10.1093/nar/gkx649.

Atomi, Haruyuki, Toshiaki Fukui, Tamotsu Kanai, Masaaki Morikawa, and Tadayuki Imanaka. 2004. 'Description of Thermococcus Kodakaraensis Sp. Nov., a Well Studied Hyperthermophilic Archaeon Previously Reported as Pyrococcus Sp. KOD1'. *Archaea (Vancouver, B.C.)* 1 (4): 263–67.

Atomi, Haruyuki, Rie Matsumi, and Tadayuki Imanaka. 2004. 'Reverse Gyrase Is Not a Prerequisite for Hyperthermophilic Life'. *Journal of Bacteriology* 186 (14): 4829–33. https://doi.org/10.1128/JB.186.14.4829-4833.2004.

Baker, Brett J., Valerie De Anda, Kiley W. Seitz, Nina Dombrowski, Alyson E. Santoro, and Karen G. Lloyd. 2020. 'Diversity, Ecology and Evolution of Archaea'. *Nature Microbiology*, May. https://doi.org/10.1038/s41564-020-0715-z.

Bates, Andrew D., and Anthony Maxwell. 2005. 'DNA Topology'. 2005. https://books.google.fr/books/about/DNA_Topology.html?id=WGBAGyzvQOUC&redir_esc=y.

Bell, S. D., C. Jaxel, M. Nadal, P. F. Kosa, and S. P. Jackson. 1998. 'Temperature, Template Topology, and Factor Requirements of Archaeal Transcription'. *Proceedings of the National Academy of Sciences of the United States of America* 95 (26): 15218–22. https://doi.org/10.1073/pnas.95.26.15218.

Bell, Stephen D., Catherine H. Botting, Benjamin N. Wardleworth, Stephen P. Jackson, and Malcolm F. White. 2002. 'The Interaction of Alba, a Conserved Archaeal Chromatin Protein, with Sir2 and Its Regulation by Acetylation'. *Science (New York, N.Y.)* 296 (5565): 148–51. https://doi.org/10.1126/science.1070506.

Benelli, Dario, and Paola Londei. 2011. 'Translation Initiation in Archaea: Conserved and Domain-Specific Features'. *Biochemical Society Transactions* 39 (1): 89–93. https://doi.org/10.1042/BST0390089.

Bliska, J. B., and N. R. Cozzarelli. 1987. 'Use of Site-Specific Recombination as a Probe of DNA Structure and Metabolism in Vivo'. *Journal of Molecular Biology* 194 (2): 205–18. https://doi.org/10.1016/0022-2836(87)90369-x.

Boles, T. C., J. H. White, and N. R. Cozzarelli. 1990. 'Structure of Plectonemically Supercoiled DNA'. *Journal of Molecular Biology* 213 (4): 931–51. https://doi.org/10.1016/S0022-2836(05)80272-4.

Borrel, Guillaume, Jean-François Brugère, Simonetta Gribaldo, Ruth A. Schmitz, and Christine Moissl-Eichinger. 2020. 'The Host-Associated Archaeome'. *Nature Reviews. Microbiology* 18 (11): 622–36. https://doi.org/10.1038/s41579-020-0407-y.

Boussau, Bastien, Samuel Blanquart, Anamaria Necsulea, Nicolas Lartillot, and Manolo Gouy. 2008. 'Parallel Adaptations to High Temperatures in the Archaean Eon'. *Nature* 456 (7224): 942–45. https://doi.org/10.1038/nature07393.

Bouthier de la Tour, C., C. Portemer, H. Kaltoum, and M. Duguet. 1998. 'Reverse Gyrase from the Hyperthermophilic Bacterium Thermotoga Maritima: Properties and Gene Structure'. *Journal of Bacteriology* 180 (2): 274–81. https://doi.org/10.1128/JB.180.2.274-281.1998.

Brochier-Armanet, Céline, Bastien Boussau, Simonetta Gribaldo, and Patrick Forterre. 2008. 'Mesophilic Crenarchaeota: Proposal for a Third Archaeal Phylum, the Thaumarchaeota'. *Nature Reviews. Microbiology* 6 (3): 245–52. https://doi.org/10.1038/nrmicro1852.

Brock, T. D., K. M. Brock, R. T. Belly, and R. L. Weiss. 1972. 'Sulfolobus: A New Genus of Sulfur-Oxidizing Bacteria Living at Low PH and High Temperature'. *Archiv Fur Mikrobiologie* 84 (1): 54–68. https://doi.org/10.1007/bf00408082.

Bult, C. J., O. White, G. J. Olsen, L. Zhou, R. D. Fleischmann, G. G. Sutton, J. A. Blake, et al. 1996. 'Complete Genome Sequence of the Methanogenic Archaeon, Methanococcus Jannaschii'. *Science (New York, N.Y.)* 273 (5278): 1058–73. https://doi.org/10.1126/science.273.5278.1058.

Bush, Natassja G., Katherine Evans-Roberts, and Anthony Maxwell. 2015. 'DNA Topoisomerases'. *EcoSal Plus* 6 (2). https://doi.org/10.1128/ecosalplus.ESP-0010-2014.

Campa, Adela G. de la, María J. Ferrándiz, Antonio J. Martín-Galiano, María T. García, and Jose M. Tirado-Vélez. 2017. 'The Transcriptome of Streptococcus Pneumoniae Induced by Local and Global Changes in Supercoiling'. *Frontiers in Microbiology* 8: 1447. https://doi.org/10.3389/fmicb.2017.01447.

Capp, Christopher, Yushen Qian, Harvey Sage, Harald Huber, and Tao-Shih Hsieh. 2010. 'Separate and Combined Biochemical Activities of the Subunits of a Naturally Split Reverse Gyrase'. *The Journal of Biological Chemistry* 285 (51): 39637–45. https://doi.org/10.1074/jbc.M110.173989.

Catchpole, Ryan J., and Patrick Forterre. 2019. 'The Evolution of Reverse Gyrase Suggests a Nonhyperthermophilic Last Universal Common Ancestor'. *Molecular Biology and Evolution* 36 (12): 2737–47. https://doi.org/10.1093/molbev/msz180.

Cavicchioli, Ricardo. 2011. 'Archaea — Timeline of the Third Domain'. *Nature Reviews Microbiology* 9 (1): 51–61. https://doi.org/10.1038/nrmicro2482.

Cavicchioli, Ricardo, Paul M. G. Curmi, Neil Saunders, and Torsten Thomas. 2003. 'Pathogenic Archaea: Do They Exist?' *BioEssays: News and Reviews in Molecular, Cellular and Developmental Biology* 25 (11): 1119–28. https://doi.org/10.1002/bies.10354.

Charbonnier, F., and P. Forterre. 1994. 'Comparison of Plasmid DNA Topology among Mesophilic and Thermophilic Eubacteria and Archaebacteria'. *Journal of Bacteriology* 176 (5): 1251–59.

Cheung, Kevin J., Vasudeo Badarinarayana, Douglas W. Selinger, Daniel Janse, and George M. Church. 2003. 'A Microarray-Based Antibiotic Screen Identifies a Regulatory Role for Supercoiling in the Osmotic Stress Response of Escherichia Coli'. *Genome Research* 13 (2): 206–15. https://doi.org/10.1101/gr.401003.

Cho, Hye Sun, Sang Sook Lee, Kwang Dong Kim, Inhwan Hwang, Jong-Seok Lim, Youn-Il Park, and Hyun-Sook Pai. 2004. 'DNA Gyrase Is Involved in Chloroplast Nucleoid Partitioning'. *The Plant Cell* 16 (10): 2665–82. https://doi.org/10.1105/tpc.104.024281.

Chu, Wai Kit, and Ian D. Hickson. 2009. 'RecQ Helicases: Multifunctional Genome Caretakers'. *Nature Reviews. Cancer* 9 (9): 644–54. https://doi.org/10.1038/nrc2682.

Cibot, Camille. n.d. 'Etude du surenroulement diffusible de l'ADN chromosomique chez la bactérie Escherichia Coli', 238.

Cockram, Charlotte, Agnès Thierry, Aurore Gorlas, Roxane Lestini, and Romain Koszul. 2021. 'Euryarchaeal Genomes Are Folded into SMC-Dependent Loops and Domains, but Lack Transcription-Mediated Compartmentalization'. *Molecular Cell* 81 (3): 459-472.e10. https://doi.org/10.1016/j.molcel.2020.12.013.

Confalonieri, F., C. Elie, M. Nadal, C. de La Tour, P. Forterre, and M. Duguet. 1993. 'Reverse Gyrase: A Helicase-like Domain and a Type I Topoisomerase in the Same Polypeptide'. *Proceedings of the National Academy of Sciences of the United States of America* 90 (10): 4753–57. https://doi.org/10.1073/pnas.90.10.4753.

Conter, Annie. 2003. 'Plasmid DNA Supercoiling and Survival in Long-Term Cultures of Escherichia Coli: Role of NaCl'. *Journal of Bacteriology* 185 (17): 5324–27. https://doi.org/10.1128/JB.185.17.5324-5327.2003.

Corte, Daniele De, Taichi Yokokawa, Marta M. Varela, Hélène Agogué, and Gerhard J. Herndl. 2009. 'Spatial Distribution of Bacteria and Archaea and Amo A Gene Copy Numbers throughout the Water Column of the Eastern Mediterranean Sea'. *The ISME Journal* 3 (2): 147–58. https://doi.org/10.1038/ismej.2008.94.

Cossu, Matteo, Catherine Badel, Ryan Catchpole, Danièle Gadelle, Evelyne Marguet, Valérie Barbe, Patrick Forterre, and Jacques Oberto. 2017. 'Flipping Chromosomes in Deep-Sea Archaea'. *PLOS Genetics* 13 (6): e1006847. https://doi.org/10.1371/journal.pgen.1006847.

Costenaro, Lionel, J. Günter Grossmann, Christine Ebel, and Anthony Maxwell. 2007. 'Modular Structure of the Full-Length DNA Gyrase B Subunit Revealed by Small-Angle X-Ray Scattering'. *Structure (London, England: 1993)* 15 (3): 329–39. https://doi.org/10.1016/j.str.2007.01.013.

Couturier, Mohea, Danièle Gadelle, Patrick Forterre, Marc Nadal, and Florence Garnier. 2020. 'The Reverse Gyrase TopR1 Is Responsible for the Homeostatic Control of DNA Supercoiling in the Hyperthermophilic Archaeon Sulfolobus Solfataricus'. *Molecular Microbiology* 113 (2): 356–68. https://doi.org/10.1111/mmi.14424.

Cremer, Thomas, and Marion Cremer. 2010. 'Chromosome Territories'. *Cold Spring Harbor Perspectives in Biology* 2 (3): a003889. https://doi.org/10.1101/cshperspect.a003889.

Da Cunha, Violette, Morgan Gaia, Daniele Gadelle, Arshan Nasir, and Patrick Forterre. 2017. 'Lokiarchaea Are Close Relatives of Euryarchaeota, Not Bridging the Gap between Prokaryotes and Eukaryotes'. *PLoS Genetics* 13 (6): e1006810. https://doi.org/10.1371/journal.pgen.1006810.

Da Cunha, Violette, Morgan Gaia, Arshan Nasir, and Patrick Forterre. 2018. 'Asgard Archaea Do Not Close the Debate about the Universal Tree of Life Topology'. *PLoS Genetics* 14 (3): e1007215. https://doi.org/10.1371/journal.pgen.1007215.

Dame, Remus T., Fatema-Zahra M. Rashid, and David C. Grainger. 2020. 'Chromosome Organization in Bacteria: Mechanistic Insights into Genome Structure and Function'. *Nature Reviews. Genetics* 21 (4): 227–42. https://doi.org/10.1038/s41576-019-0185-4.

Dar, Mohd Ashraf, Atul Sharma, Neelima Mondal, and Suman Kumar Dhar. 2007. 'Molecular Cloning of Apicoplast-Targeted Plasmodium Falciparum DNA Gyrase Genes: Unique Intrinsic ATPase Activity and ATP-Independent Dimerization of PfGyrB Subunit'. *Eukaryotic Cell* 6 (3): 398–412. https://doi.org/10.1128/EC.00357-06.

Darwin, Charles. 1859. 'On the Origin of Species by Charles Darwin'.

De Rosa, Mario, Agata Gambacorta, Barbara Nicolaus, Salvatore Sodano, and J.D. Bu'Lock. 1980. 'Structural Regularities in Tetraether Lipids of Caldariella and Their Biosynthetic and Phyletic Implications'. *Phytochemistry* 19 (5): 833–36. https://doi.org/10.1016/0031-9422(80)85121-1.

Déclais, A. C., J. Marsault, F. Confalonieri, C. B. de La Tour, and M. Duguet. 2000. 'Reverse Gyrase, the Two Domains Intimately Cooperate to Promote Positive Supercoiling'. *The Journal of Biological Chemistry* 275 (26): 19498–504. https://doi.org/10.1074/jbc.M910091199.

DeLong, E. F. 1992. 'Archaea in Coastal Marine Environments'. *Proceedings of the National Academy of Sciences of the United States of America* 89 (12): 5685–89. https://doi.org/10.1073/pnas.89.12.5685.

DeLong, E. F. 1998. 'Everything in Moderation: Archaea as "Non-Extremophiles"'. *Current Opinion in Genetics & Development* 8 (6): 649–54. https://doi.org/10.1016/s0959-437x(98)80032-4.

Dh, Parks, Chuvochina M, Waite Dw, Rinke C, Skarshewski A, Chaumeil Pa, and Hugenholtz P. 2018. 'A Standardized Bacterial Taxonomy Based on Genome Phylogeny Substantially Revises the Tree of Life'. Nature Biotechnology. Nat Biotechnol. November 2018. https://doi.org/10.1038/nbt.4229.

Dillon, Shane C., and Charles J. Dorman. 2010. 'Bacterial Nucleoid-Associated Proteins, Nucleoid Structure and Gene Expression'. *Nature Reviews. Microbiology* 8 (3): 185–95. https://doi.org/10.1038/nrmicro2261.

Dorman, C. J., G. C. Barr, N. Ni Bhriain, and C. F. Higgins. 1988. 'DNA Supercoiling and the Anaerobic and Growth Phase Regulation of TonB Gene Expression'. *Journal of Bacteriology* 170 (6): 2816–26.

Dorman, C. J., N. Ni Bhriain, and C. F. Higgins. 1990. 'DNA Supercoiling and Environmental Regulation of Virulence Gene Expression in Shigella Flexneri'. *Nature* 344 (6268): 789–92. https://doi.org/10.1038/344789a0.

Dorman, Charles J. 2006. 'DNA Supercoiling and Bacterial Gene Expression'. *Science Progress* 89 (Pt 3-4): 151–66.

Driessen, Rosalie P. C., and Remus Th Dame. 2011. 'Nucleoid-Associated Proteins in Crenarchaea'. *Biochemical Society Transactions* 39 (1): 116–21. https://doi.org/10.1042/BST0390116.

Drlica, K. 1992. 'Control of Bacterial DNA Supercoiling'. *Molecular Microbiology* 6 (4): 425–33. https://doi.org/10.1111/j.1365-2958.1992.tb01486.x.

Drolet, Marc. 2006. 'Growth Inhibition Mediated by Excess Negative Supercoiling: The Interplay between Transcription Elongation, R-Loop Formation and DNA Topology'. *Molecular Microbiology* 59 (3): 723–30. https://doi.org/10.1111/j.1365-2958.2005.05006.x.

Kouzine, F., Baranello, L., Levens, D., 2018. *DNA Topoisomerases: Methods and Protocols*. Methods in Molecular Biology. Humana Press. //www.springer.com/us/book/9781493974580.

Efremov, Artem K., Yuanyuan Qu, Hugo Maruyama, Ci J. Lim, Kunio Takeyasu, and Jie Yan. 2015. 'Transcriptional Repressor TrmBL2 from Thermococcus Kodakarensis Forms Filamentous Nucleoprotein Structures and Competes with Histones for DNA Binding in a Salt- and DNA Supercoiling-Dependent Manner'. *The Journal of Biological Chemistry* 290 (25): 15770–84. https://doi.org/10.1074/jbc.M114.626705.

Ettema, Thijs J. G., Ann-Christin Lindås, and Rolf Bernander. 2011. 'An Actin-Based Cytoskeleton in Archaea'. *Molecular Microbiology* 80 (4): 1052–61. https://doi.org/10.1111/j.1365-2958.2011.07635.x.

Evans, Paul N., Joel A. Boyd, Andy O. Leu, Ben J. Woodcroft, Donovan H. Parks, Philip Hugenholtz, and Gene W. Tyson. 2019. 'An Evolving View of Methane Metabolism in the Archaea'. *Nature Reviews Microbiology* 17 (4): 219–32. https://doi.org/10.1038/s41579-018-0136-7.

Evans, Paul N., Donovan H. Parks, Grayson L. Chadwick, Steven J. Robbins, Victoria J. Orphan, Suzanne D. Golding, and Gene W. Tyson. 2015. 'Methane Metabolism in the Archaeal Phylum Bathyarchaeota Revealed by Genome-Centric Metagenomics'. *Science (New York, N.Y.)* 350 (6259): 434–38. https://doi.org/10.1126/science.aac7745.

Ferrándiz, María-José, Antonio J. Martín-Galiano, Cristina Arnanz, Isabel Camacho-Soguero, José-Manuel Tirado-Vélez, and Adela G. de la Campa. 2016. 'An Increase in Negative Supercoiling in Bacteria Reveals Topology-Reacting Gene Clusters and a Homeostatic Response Mediated by the DNA Topoisomerase I Gene'. *Nucleic Acids Research* 44 (15): 7292–7303. https://doi.org/10.1093/nar/gkw602.

Forterre, P., F. Confalonieri, and S. Knapp. 1999. 'Identification of the Gene Encoding Archeal-Specific DNA-Binding Proteins of the Sac10b Family'. *Molecular Microbiology* 32 (3): 669–70. https://doi.org/10.1046/j.1365-2958.1999.01366.x.

Forterre, Patrick. 2002. 'A Hot Story from Comparative Genomics: Reverse Gyrase Is the Only Hyperthermophile-Specific Protein'. *Trends in Genetics: TIG* 18 (5): 236–37.

Forterre, Patrick, and Danièle Gadelle. 2009. 'Phylogenomics of DNA Topoisomerases: Their Origin and Putative Roles in the Emergence of Modern Organisms'. *Nucleic Acids Research* 37 (3): 679–92. https://doi.org/10.1093/nar/gkp032.

Forterre, Patrick, Simonetta Gribaldo, Danièle Gadelle, and Marie-Claude Serre. 2007. 'Origin and Evolution of DNA Topoisomerases'. *Biochimie*, DNA Topology, 89 (4): 427–46. https://doi.org/10.1016/j.biochi.2006.12.009.

Fournier, Gregory P., and Anthony M. Poole. 2018. 'A Briefly Argued Case That Asgard Archaea Are Part of the Eukaryote Tree'. *Frontiers in Microbiology* 9: 1896. https://doi.org/10.3389/fmicb.2018.01896.

Fox, G. E., L. J. Magrum, W. E. Balch, R. S. Wolfe, and C. R. Woese. 1977. 'Classification of Methanogenic Bacteria by 16S Ribosomal RNA Characterization'. *Proceedings of the National Academy of Sciences of the United States of America* 74 (10): 4537–41. https://doi.org/10.1073/pnas.74.10.4537.

Franzmann, P. D., Y. Liu, D. L. Balkwill, H. C. Aldrich, E. Conway de Macario, and D. R. Boone. 1997. 'Methanogenium Frigidum Sp. Nov., a Psychrophilic, H2-Using Methanogen from Ace Lake, Antarctica'. *International Journal of Systematic Bacteriology* 47 (4): 1068–72. https://doi.org/10.1099/00207713-47-4-1068.

Franzmann, P. D., N. Springer, W. Ludwig, E. Conway De Macario, and M. Rohde. 1992. 'A Methanogenic Archaeon from Ace Lake, Antarctica: Methanococcoides Burtonii Sp. Nov.' *Systematic and Applied Microbiology* 15 (4): 573–81. https://doi.org/10.1016/S0723-2020(11)80117-7.

Fuhrman, J. A., K. McCallum, and A. A. Davis. 1992. 'Novel Major Archaebacterial Group from Marine Plankton'. *Nature* 356 (6365): 148–49. https://doi.org/10.1038/356148a0.

Fütterer, O., A. Angelov, H. Liesegang, G. Gottschalk, C. Schleper, B. Schepers, C. Dock, G. Antranikian, and W. Liebl. 2004. 'Genome Sequence of Picrophilus Torridus and Its Implications for Life around PH 0'. *Proceedings of the National Academy of Sciences of the United States of America* 101 (24): 9091–96. https://doi.org/10.1073/pnas.0401356101.

Fyodorov, Dmitry V., Bing-Rui Zhou, Arthur I. Skoultchi, and Yawen Bai. 2018. 'Emerging Roles of Linker Histones in Regulating Chromatin Structure and Function'. *Nature Reviews. Molecular Cell Biology* 19 (3): 192–206. https://doi.org/10.1038/nrm.2017.94.

Ganguly, Agneyo, Yoandris del Toro Duany, and Dagmar Klostermeier. 2013. 'Reverse Gyrase Transiently Unwinds Double-Stranded DNA in an ATP-Dependent Reaction'. *Journal of Molecular Biology* 425 (1): 32–40. https://doi.org/10.1016/j.jmb.2012.10.016.

Garnier, Florence, Mohea Couturier, Hélène Débat, and Marc Nadal. 2021. 'Archaea: A Gold Mine for Topoisomerase Diversity'. *Frontiers in Microbiology* 12: 661411. https://doi.org/10.3389/fmicb.2021.661411.

Gellert, M., K. Mizuuchi, M. H. O'Dea, T. Itoh, and J. I. Tomizawa. 1977. 'Nalidixic Acid Resistance: A Second Genetic Character Involved in DNA Gyrase Activity'. *Proceedings of the National Academy of Sciences of the United States of America* 74 (11): 4772–76. https://doi.org/10.1073/pnas.74.11.4772.

Gellert, M., K. Mizuuchi, M. H. O'Dea, and H. A. Nash. 1976. 'DNA Gyrase: An Enzyme That Introduces Superhelical Turns into DNA'. *Proceedings of the National Academy of Sciences of the United States of America* 73 (11): 3872–76. https://doi.org/10.1073/pnas.73.11.3872.

Gellert, M., M. H. O'Dea, T. Itoh, and J. Tomizawa. 1976. 'Novobiocin and Coumermycin Inhibit DNA Supercoiling Catalyzed by DNA Gyrase'. *Proceedings of the National Academy of Sciences of the United States of America* 73 (12): 4474–78. https://doi.org/10.1073/pnas.73.12.4474.

Green, G. R., D. G. Searcy, and R. J. DeLange. 1983. 'Histone-like Protein in the Archaebacterium Sulfolobus Acidocaldarius'. *Biochimica Et Biophysica Acta* 741 (2): 251–57. https://doi.org/10.1016/0167-4781(83)90066-0.

Grogan, D. W. 1998. 'Hyperthermophiles and the Problem of DNA Instability'. *Molecular Microbiology* 28 (6): 1043–49. https://doi.org/10.1046/j.1365-2958.1998.00853.x.
Grogan, D. W. 2000. 'The Question of DNA Repair in Hyperthermophilic Archaea'. *Trends in Microbiology* 8 (4): 180–85. https://doi.org/10.1016/s0966-842x(00)01729-7.
Großkopf, Regine, Stephan Stubner, and Werner Liesack. 1998. 'Novel Euryarchaeotal Lineages Detected on Rice Roots and in the Anoxic Bulk Soil of Flooded Rice Microcosms'. *Applied and Environmental Microbiology* 64 (12): 4983–89.
Groussin, Mathieu, and Manolo Gouy. 2011. 'Adaptation to Environmental Temperature Is a Major Determinant of Molecular Evolutionary Rates in Archaea'. *Molecular Biology and Evolution* 28 (9): 2661–74. https://doi.org/10.1093/molbev/msr098.
Guipaud, O., E. Marguet, K. M. Noll, C. B. de la Tour, and P. Forterre. 1997. 'Both DNA Gyrase and Reverse Gyrase Are Present in the Hyperthermophilic Bacterium Thermotoga Maritima'. *Proceedings of the National Academy of Sciences of the United States of America* 94 (20): 10606–11.
Han, Wenyuan, Xu Feng, and Qunxin She. 2017. 'Reverse Gyrase Functions in Genome Integrity Maintenance by Protecting DNA Breaks In Vivo'. *International Journal of Molecular Sciences* 18 (7): E1340. https://doi.org/10.3390/ijms18071340.
Harish, Ajith, and Charles G. Kurland. 2017a. 'Akaryotes and Eukaryotes Are Independent Descendants of a Universal Common Ancestor'. *Biochimie* 138 (July): 168–83. https://doi.org/10.1016/j.biochi.2017.04.013.
Harish, Ajith, and Charles G. Kurland. 2017b. 'Empirical Genome Evolution Models Root the Tree of Life'. *Biochimie* 138 (July): 137–55. https://doi.org/10.1016/j.biochi.2017.04.014.
Henneman, Bram, Thomas B. Brouwer, Amanda M. Erkelens, Gert-Jan Kuijntjes, Clara van Emmerik, Ramon A. van der Valk, Monika Timmer, et al. 2021. 'Mechanical and Structural Properties of Archaeal Hypernucleosomes'. *Nucleic Acids Research* 49 (8): 4338–49. https://doi.org/10.1093/nar/gkaa1196.
Henneman, Bram, Clara van Emmerik, Hugo van Ingen, and Remus T. Dame. 2018. 'Structure and Function of Archaeal Histones'. *PLOS Genetics* 14 (9): e1007582. https://doi.org/10.1371/journal.pgen.1007582.
Hethke, C., A. Bergerat, W. Hausner, P. Forterre, and M. Thomm. 1999. 'Cell-Free Transcription at 95 Degrees: Thermostability of Transcriptional Components and DNA Topology Requirements of Pyrococcus Transcription'. *Genetics* 152 (4): 1325–33. https://doi.org/10.1093/genetics/152.4.1325.
Higgins, C. F., C. J. Dorman, D. A. Stirling, L. Waddell, I. R. Booth, G. May, and E. Bremer. 1988. 'A Physiological Role for DNA Supercoiling in the Osmotic Regulation of Gene Expression in S. Typhimurium and E. Coli'. *Cell* 52 (4): 569–84. https://doi.org/10.1016/0092-8674(88)90470-9.
Higgins, N. P., C. L. Peebles, A. Sugino, and N. R. Cozzarelli. 1978. 'Purification of Subunits of Escherichia Coli DNA Gyrase and Reconstitution of Enzymatic Activity'. *Proceedings of the National Academy of Sciences of the United States of America* 75 (4): 1773–77. https://doi.org/10.1073/pnas.75.4.1773.
Hileman, Travis H., and Thomas J. Santangelo. 2012. 'Genetics Techniques for Thermococcus Kodakarensis'. *Frontiers in Microbiology* 3: 195. https://doi.org/10.3389/fmicb.2012.00195.
Hirata, Akira, Brianna J. Klein, and Katsuhiko S. Murakami. 2008. 'The X-Ray Crystal Structure of RNA Polymerase from Archaea'. *Nature* 451 (7180): 851–54. https://doi.org/10.1038/nature06530.
Hirsch, Jana, and Dagmar Klostermeier. 2021. 'What Makes a Type IIA Topoisomerase a Gyrase or a Topo IV?' *Nucleic Acids Research* 49 (11): 6027–42. https://doi.org/10.1093/nar/gkab270.
Hocher, Antoine, Guillaume Borrel, Khaled Fadhlaoui, Jean-François Brugère, Simonetta Gribaldo, and Tobias Warnecke. 2021. 'Growth Temperature Is the Principal Driver of Chromatinization in Archaea'. Preprint. Evolutionary Biology. https://doi.org/10.1101/2021.07.08.451601.

Hocher, Antoine, Maria Rojec, Jacob B. Swadling, Alexander Esin, and Tobias Warnecke. 2019. 'The DNA-Binding Protein HTa from Thermoplasma Acidophilum Is an Archaeal Histone Analog'. *ELife* 8 (November). https://doi.org/10.7554/eLife.52542.

Holmes, M. L., and M. L. Dyall-Smith. 1991. 'Mutations in DNA Gyrase Result in Novobiocin Resistance in Halophilic Archaebacteria'. *Journal of Bacteriology* 173 (2): 642–48.

Hołówka, Joanna, and Jolanta Zakrzewska-Czerwińska. 2020. 'Nucleoid Associated Proteins: The Small Organizers That Help to Cope With Stress'. *Frontiers in Microbiology* 0. https://doi.org/10.3389/fmicb.2020.00590.

Huber, R., T. Langworthy, H. König, M. Thomm, C. Woese, U. Sleytr, and K. Stetter. 2004. 'Thermotoga Maritima Sp. Nov. Represents a New Genus of Unique Extremely Thermophilic Eubacteria Growing up to 90°C'. *Archives of Microbiology*. https://doi.org/10.1007/BF00409880.

Huet, J, R Schnabel, A Sentenac, and W Zillig. 1983. 'Archaebacteria and Eukaryotes Possess DNA-Dependent RNA Polymerases of a Common Type.' *The EMBO Journal* 2 (8): 1291–94.

Imachi, Hiroyuki, Masaru K. Nobu, Nozomi Nakahara, Yuki Morono, Miyuki Ogawara, Yoshihiro Takaki, Yoshinori Takano, et al. 2020. 'Isolation of an Archaeon at the Prokaryote-Eukaryote Interface'. *Nature* 577 (7791): 519–25. https://doi.org/10.1038/s41586-019-1916-6.

Ishino, Yoshizumi, and Issay Narumi. 2015. 'DNA Repair in Hyperthermophilic and Hyperradioresistant Microorganisms'. *Current Opinion in Microbiology* 25 (June): 103–12. https://doi.org/10.1016/j.mib.2015.05.010.

Janssen, Aniek, Serafin U. Colmenares, and Gary H. Karpen. 2018. 'Heterochromatin: Guardian of the Genome'. *Annual Review of Cell and Developmental Biology* 34: 265–88. https://doi.org/10.1146/annurev-cellbio-100617-062653.

Kachlany, Scott C., Paul J. Planet, Mrinal K. Bhattacharjee, Evyenia Kollia, Rob DeSalle, Daniel H. Fine, and David H. Figurski. 2000. 'Nonspecific Adherence by Actinobacillus Actinomycetemcomitans Requires Genes Widespread in Bacteria and Archaea'. *Journal of Bacteriology* 182 (21): 6169–76.

Kampmann, Martin, and Daniela Stock. 2004. 'Reverse Gyrase Has Heat-Protective DNA Chaperone Activity Independent of Supercoiling'. *Nucleic Acids Research* 32 (12): 3537–45. https://doi.org/10.1093/nar/gkh683.

Karem, K., and J. W. Foster. 1993. 'The Influence of DNA Topology on the Environmental Regulation of a PH-Regulated Locus in Salmonella Typhimurium'. *Molecular Microbiology* 10 (1): 75–86. https://doi.org/10.1111/j.1365-2958.1993.tb00905.x.

Karner, M. B., E. F. DeLong, and D. M. Karl. 2001. 'Archaeal Dominance in the Mesopelagic Zone of the Pacific Ocean'. *Nature* 409 (6819): 507–10. https://doi.org/10.1038/35054051.

Kempfer, Rieke, and Ana Pombo. 2019. 'Methods for Mapping 3D Chromosome Architecture'. *Nature Reviews Genetics*, December, 1–20. https://doi.org/10.1038/s41576-019-0195-2.

Kikuchi, A., and K. Asai. 1984. 'Reverse Gyrase--a Topoisomerase Which Introduces Positive Superhelical Turns into DNA'. *Nature* 309 (5970): 677–81.

Klenk, H. P., R. A. Clayton, J. F. Tomb, O. White, K. E. Nelson, K. A. Ketchum, R. J. Dodson, et al. 1997. 'The Complete Genome Sequence of the Hyperthermophilic, Sulphate-Reducing Archaeon Archaeoglobus Fulgidus'. *Nature* 390 (6658): 364–70. https://doi.org/10.1038/37052.

Klevan, L., and J. C. Wang. 1980. 'Deoxyribonucleic Acid Gyrase-Deoxyribonucleic Acid Complex Containing 140 Base Pairs of Deoxyribonucleic Acid and an Alpha 2 Beta 2 Protein Core'. *Biochemistry* 19 (23): 5229–34. https://doi.org/10.1021/bi00564a012.

Koga, Yosuke. 2012. 'Thermal Adaptation of the Archaeal and Bacterial Lipid Membranes'. *Archaea (Vancouver, B.C.)* 2012: 789652. https://doi.org/10.1155/2012/789652.

Koonin, Eugene V. 2009. 'Evolution of Genome Architecture'. *The International Journal of Biochemistry & Cell Biology* 41 (2): 298–306. https://doi.org/10.1016/j.biocel.2008.09.015.

Koskinen, Kaisa, Manuela R. Pausan, Alexandra K. Perras, Michael Beck, Corinna Bang, Maximilian Mora, Anke Schilhabel, Ruth Schmitz, and Christine Moissl-Eichinger. 2017. 'First Insights into

the Diverse Human Archaeome: Specific Detection of Archaea in the Gastrointestinal Tract, Lung, and Nose and on Skin'. *MBio* 8 (6). https://doi.org/10.1128/mBio.00824-17.

Koster, Daniel A., Aurélien Crut, Stewart Shuman, Mary-Ann Bjornsti, and Nynke H. Dekker. 2010. 'Cellular Strategies for Regulating DNA Supercoiling: A Single-Molecule Perspective'. *Cell* 142 (4): 519–30. https://doi.org/10.1016/j.cell.2010.08.001.

Kouzine, Fedor, Ashutosh Gupta, Laura Baranello, Damian Wojtowicz, Khadija Ben-Aissa, Juhong Liu, Teresa M Przytycka, and David Levens. 2013. 'Transcription-Dependent Dynamic Supercoiling Is a Short-Range Genomic Force'. *Nature Structural & Molecular Biology* 20 (3): 396–403. https://doi.org/10.1038/nsmb.2517.

Kozyavkin, S. A., R. Krah, M. Gellert, K. O. Stetter, J. A. Lake, and A. I. Slesarev. 1994. 'A Reverse Gyrase with an Unusual Structure. A Type I DNA Topoisomerase from the Hyperthermophile Methanopyrus Kandleri Is a Two-Subunit Protein'. *The Journal of Biological Chemistry* 269 (15): 11081–89.

Kramlinger, Valerie M., and Hiroshi Hiasa. 2006. 'The "GyrA-Box" Is Required for the Ability of DNA Gyrase to Wrap DNA and Catalyze the Supercoiling Reaction'. *The Journal of Biological Chemistry* 281 (6): 3738–42. https://doi.org/10.1074/jbc.M511160200.

Krassovsky, Kristina, Rajarshi P. Ghosh, and Barbara J. Meyer. 2021. 'Genome-Wide Profiling Reveals Functional Interplay of DNA Sequence Composition, Transcriptional Activity, and Nucleosome Positioning in Driving DNA Supercoiling and Helix Destabilization in C. Elegans'. *Genome Research*, June. https://doi.org/10.1101/gr.270082.120.

Kreuzer, K. N., and N. R. Cozzarelli. 1980. 'Formation and Resolution of DNA Catenanes by DNA Gyrase'. *Cell* 20 (1): 245–54. https://doi.org/10.1016/0092-8674(80)90252-4.

Lake, J. A., E. Henderson, M. Oakes, and M. W. Clark. 1984. 'Eocytes: A New Ribosome Structure Indicates a Kingdom with a Close Relationship to Eukaryotes'. *Proceedings of the National Academy of Sciences* 81 (12): 3786–90. https://doi.org/10.1073/pnas.81.12.3786.

Lal, Avantika, Amlanjyoti Dhar, Andrei Trostel, Fedor Kouzine, Aswin S. N. Seshasayee, and Sankar Adhya. 2016. 'Genome Scale Patterns of Supercoiling in a Bacterial Chromosome'. *Nature Communications* 7 (March). https://doi.org/10.1038/ncomms11055.

Lane, Nick, and William Martin. 2010. 'The Energetics of Genome Complexity'. *Nature* 467 (7318): 929–34. https://doi.org/10.1038/nature09486.

Langworthy, Thomas A., William R. Mayberry, and Paul F. Smith. 1974. 'Long-Chain Glycerol Diether and Polyol Dialkyl Glycerol Triether Lipids of Sulfolobus Acidocaldarius'. *Journal of Bacteriology* 119 (1): 106–16.

Lee, Sung-Jae, Melanie Surma, Sabine Seitz, Winfried Hausner, Michael Thomm, and Winfried Boos. 2007. 'Characterization of the TrmB-like Protein, PF0124, a TGM-Recognizing Global Transcriptional Regulator of the Hyperthermophilic Archaeon Pyrococcus Furiosus'. *Molecular Microbiology* 65 (2): 305–18. https://doi.org/10.1111/j.1365-2958.2007.05780.x.

Lindahl, T. 1993. 'Instability and Decay of the Primary Structure of DNA'. *Nature* 362 (6422): 709–15. https://doi.org/10.1038/362709a0.

Lindås, Ann-Christin, Erik A. Karlsson, Maria T. Lindgren, Thijs J. G. Ettema, and Rolf Bernander. 2008. 'A Unique Cell Division Machinery in the Archaea'. *Proceedings of the National Academy of Sciences of the United States of America* 105 (48): 18942–46. https://doi.org/10.1073/pnas.0809467105.

Lioy, Virginia S., Axel Cournac, Martial Marbouty, Stéphane Duigou, Julien Mozziconacci, Olivier Espéli, Frédéric Boccard, and Romain Koszul. 2018. 'Multiscale Structuring of the E. Coli Chromosome by Nucleoid-Associated and Condensin Proteins'. *Cell* 172 (4): 771-783.e18. https://doi.org/10.1016/j.cell.2017.12.027.

Lipscomb, Gina L., Elin M. Hahn, Alexander T. Crowley, and Michael W. W. Adams. 2017. 'Reverse Gyrase Is Essential for Microbial Growth at 95 °C'. *Extremophiles: Life Under Extreme Conditions* 21 (3): 603–8. https://doi.org/10.1007/s00792-017-0929-z.

Liu, L. F., C. C. Liu, and B. M. Alberts. 1980. 'Type II DNA Topoisomerases: Enzymes That Can Unknot a Topologically Knotted DNA Molecule via a Reversible Double-Strand Break'. *Cell* 19 (3): 697–707. https://doi.org/10.1016/s0092-8674(80)80046-8.

Liu, L. F., and J. C. Wang. 1987. 'Supercoiling of the DNA Template during Transcription'. *Proceedings of the National Academy of Sciences of the United States of America* 84 (20): 7024–27.

López-García, P. 1999. 'DNA Supercoiling and Temperature Adaptation: A Clue to Early Diversification of Life?' *Journal of Molecular Evolution* 49 (4): 439–52. https://doi.org/10.1007/pl00006567.

López-García, P., and P. Forterre. 1997. 'DNA Topology in Hyperthermophilic Archaea: Reference States and Their Variation with Growth Phase, Growth Temperature, and Temperature Stresses'. *Molecular Microbiology* 23 (6): 1267–79.

López-García, P., and P. Forterre. 1999. 'Control of DNA Topology during Thermal Stress in Hyperthermophilic Archaea: DNA Topoisomerase Levels, Activities and Induced Thermotolerance during Heat and Cold Shock in Sulfolobus'. *Molecular Microbiology* 33 (4): 766–77.

López-García, Purificación, Patrick Forterre, John van der Oost, and Gaël Erauso. 2000. 'Plasmid PGS5 from the Hyperthermophilic Archaeon Archaeoglobus Profundus Is Negatively Supercoiled'. *Journal of Bacteriology* 182 (17): 4998–5000.

López-García, Purificación, Yvan Zivanovic, Philippe Deschamps, and David Moreira. 2015. 'Bacterial Gene Import and Mesophilic Adaptation in Archaea'. *Nature Reviews. Microbiology* 13 (7): 447–56. https://doi.org/10.1038/nrmicro3485.

Luger, K., A. W. Mäder, R. K. Richmond, D. F. Sargent, and T. J. Richmond. 1997. 'Crystal Structure of the Nucleosome Core Particle at 2.8 A Resolution'. *Nature* 389 (6648): 251–60. https://doi.org/10.1038/38444.

Lulchev, Pavel, and Dagmar Klostermeier. 2014. 'Reverse Gyrase—Recent Advances and Current Mechanistic Understanding of Positive DNA Supercoiling'. *Nucleic Acids Research* 42 (13): 8200–8213. https://doi.org/10.1093/nar/gku589.

Makarova, Kira S., Eugene V. Koonin, and Sonja-Verena Albers. 2016. 'Diversity and Evolution of Type IV Pili Systems in Archaea'. *Frontiers in Microbiology* 7. https://doi.org/10.3389/fmicb.2016.00667.

Makarova, Kira S., Yuri I. Wolf, and Eugene V. Koonin. 2003. 'Potential Genomic Determinants of Hyperthermophily'. *Trends in Genetics* 19 (4): 172–76. https://doi.org/10.1016/S0168-9525(03)00047-7.

Marguet, E., and P. Forterre. 1994. 'DNA Stability at Temperatures Typical for Hyperthermophiles'. *Nucleic Acids Research* 22 (9): 1681–86. https://doi.org/10.1093/nar/22.9.1681.

Marshall, C. L., and A. D. Brown. 1968. 'The Membrane Lipids of Halobacterium Halobium'. *The Biochemical Journal* 110 (3): 441–48. https://doi.org/10.1042/bj1100441.

Marteinsson, V. T., J. L. Birrien, A. L. Reysenbach, M. Vernet, D. Marie, A. Gambacorta, P. Messner, U. B. Sleytr, and D. Prieur. 1999. 'Thermococcus Barophilus Sp. Nov., a New Barophilic and Hyperthermophilic Archaeon Isolated under High Hydrostatic Pressure from a Deep-Sea Hydrothermal Vent'. *International Journal of Systematic Bacteriology* 49 Pt 2 (April): 351–59. https://doi.org/10.1099/00207713-49-2-351.

Maruyama, Hugo, Janet C Harwood, Karen M Moore, Konrad Paszkiewicz, Samuel C Durley, Hisanori Fukushima, Haruyuki Atomi, Kunio Takeyasu, and Nicholas A Kent. 2013. 'An Alternative Beads-on-a-String Chromatin Architecture in Thermococcus Kodakarensis'. *EMBO Reports* 14 (8): 711–17. https://doi.org/10.1038/embor.2013.94.

Maruyama, Hugo, Minsang Shin, Toshiyuki Oda, Rie Matsumi, Ryosuke L. Ohniwa, Takehiko Itoh, Katsuhiko Shirahige, et al. 2011. 'Histone and TK0471/TrmBL2 Form a Novel Heterogeneous Genome Architecture in the Hyperthermophilic Archaeon Thermococcus Kodakarensis'. *Molecular Biology of the Cell* 22 (3): 386–98. https://doi.org/10.1091/mbc.e10-08-0668.

Mattiroli, Francesca, Sudipta Bhattacharyya, Pamela N. Dyer, Alison E. White, Kathleen Sandman, Brett W. Burkhart, Kyle R. Byrne, et al. 2017. 'Structure of Histone-Based Chromatin in

Archaea'. *Science (New York, N.Y.)* 357 (6351): 609–12. https://doi.org/10.1126/science.aaj1849.
Menzel, Rolf, and Martin Gellert. 1983. 'Regulation of the Genes for E. Coli DNA Gyrase: Homeostatic Control of DNA Supercoiling'. *Cell* 34 (1): 105–13. https://doi.org/10.1016/0092-8674(83)90140-X.
Meselson, M., and F. W. Stahl. 1958. 'THE REPLICATION OF DNA IN ESCHERICHIA COLI'. *Proceedings of the National Academy of Sciences of the United States of America* 44 (7): 671–82.
Mizuuchi, K., L. M. Fisher, M. H. O'Dea, and M. Gellert. 1980. 'DNA Gyrase Action Involves the Introduction of Transient Double-Strand Breaks into DNA'. *Proceedings of the National Academy of Sciences of the United States of America* 77 (4): 1847–51. https://doi.org/10.1073/pnas.77.4.1847.
Morimoto, Nanako, Wakao Fukuda, Nanami Nakajima, Takeaki Masuda, Yusuke Terui, Tamotsu Kanai, Tairo Oshima, Tadayuki Imanaka, and Shinsuke Fujiwara. 2010. 'Dual Biosynthesis Pathway for Longer-Chain Polyamines in the Hyperthermophilic Archaeon Thermococcus Kodakarensis'. *Journal of Bacteriology* 192 (19): 4991–5001. https://doi.org/10.1128/JB.00279-10.
Mullis, Megan M., Ian M. Rambo, Brett J. Baker, and Brandi Kiel Reese. 2019. 'Diversity, Ecology, and Prevalence of Antimicrobials in Nature'. *Frontiers in Microbiology* 10: 2518. https://doi.org/10.3389/fmicb.2019.02518.
Musgrave, D., P. Forterre, and A. Slesarev. 2000. 'Negative Constrained DNA Supercoiling in Archaeal Nucleosomes'. *Molecular Microbiology* 35 (2): 341–49. https://doi.org/10.1046/j.1365-2958.2000.01689.x.
Musgrave, D. R., K. M. Sandman, and J. N. Reeve. 1991. 'DNA Binding by the Archaeal Histone HMf Results in Positive Supercoiling'. *Proceedings of the National Academy of Sciences of the United States of America* 88 (23): 10397–401. https://doi.org/10.1073/pnas.88.23.10397.
Napoli, Alessandra, Anna Valenti, Vincenzo Salerno, Marc Nadal, Florence Garnier, Mosè Rossi, and Maria Ciaramella. 2004. 'Reverse Gyrase Recruitment to DNA after UV Light Irradiation in Sulfolobus Solfataricus'. *The Journal of Biological Chemistry* 279 (32): 33192–98. https://doi.org/10.1074/jbc.M402619200.
Napoli, Alessandra, Anna Valenti, Vincenzo Salerno, Marc Nadal, Florence Garnier, Mosè Rossi, and Maria Ciaramella. 2005. 'Functional Interaction of Reverse Gyrase with Single-Strand Binding Protein of the Archaeon Sulfolobus'. *Nucleic Acids Research* 33 (2): 564–76. https://doi.org/10.1093/nar/gki202.
Niki, H., Y. Yamaichi, and S. Hiraga. 2000. 'Dynamic Organization of Chromosomal DNA in Escherichia Coli'. *Genes & Development* 14 (2): 212–23.
Niki, Hironori, and Sota Hiraga. 1998. 'Polar Localization of the Replication Origin and Terminus in Escherichia Coli Nucleoids during Chromosome Partitioning'. *Genes & Development* 12 (7): 1036–45.
O'Connor, E. M., and R. F. Shand. 2002. 'Halocins and Sulfolobicins: The Emerging Story of Archaeal Protein and Peptide Antibiotics'. *Journal of Industrial Microbiology & Biotechnology* 28 (1): 23–31. https://doi.org/10.1038/sj/jim/7000190.
Offre, Pierre, Anja Spang, and Christa Schleper. 2013. 'Archaea in Biogeochemical Cycles'. *Annual Review of Microbiology* 67: 437–57. https://doi.org/10.1146/annurev-micro-092412-155614.
Oren, Aharon. 2002. 'Molecular Ecology of Extremely Halophilic Archaea and Bacteria'. *FEMS Microbiology Ecology* 39 (1): 1–7. https://doi.org/10.1111/j.1574-6941.2002.tb00900.x.
Parks, Donovan H., Maria Chuvochina, Pierre-Alain Chaumeil, Christian Rinke, Aaron J. Mussig, and Philip Hugenholtz. 2020. 'A Complete Domain-to-Species Taxonomy for Bacteria and Archaea'. *Nature Biotechnology*, April, 1–8. https://doi.org/10.1038/s41587-020-0501-8.
Pereira, S. L., R. A. Grayling, R. Lurz, and J. N. Reeve. 1997. 'Archaeal Nucleosomes'. *Proceedings of the National Academy of Sciences of the United States of America* 94 (23): 12633–37. https://doi.org/10.1073/pnas.94.23.12633.

Perugino, Giuseppe, Anna Valenti, Anna D'amaro, Mosè Rossi, and Maria Ciaramella. 2009. 'Reverse Gyrase and Genome Stability in Hyperthermophilic Organisms'. *Biochemical Society Transactions* 37 (Pt 1): 69–73. https://doi.org/10.1042/BST0370069.

Peter, Brian J., Javier Arsuaga, Adam M. Breier, Arkady B. Khodursky, Patrick O. Brown, and Nicholas R. Cozzarelli. 2004. 'Genomic Transcriptional Response to Loss of Chromosomal Supercoiling in Escherichia Coli'. *Genome Biology* 5 (11): R87. https://doi.org/10.1186/gb-2004-5-11-r87.

Petitjean, Céline, Philippe Deschamps, Purificación López-García, and David Moreira. 2014. 'Rooting the Domain Archaea by Phylogenomic Analysis Supports the Foundation of the New Kingdom Proteoarchaeota'. *Genome Biology and Evolution* 7 (1): 191–204. https://doi.org/10.1093/gbe/evu274.

Postow, L., N. J. Crisona, B. J. Peter, C. D. Hardy, and N. R. Cozzarelli. 2001. 'Topological Challenges to DNA Replication: Conformations at the Fork'. *Proceedings of the National Academy of Sciences of the United States of America* 98 (15): 8219–26. https://doi.org/10.1073/pnas.111006998.

Pruss, G. J., S. H. Manes, and K. Drlica. 1982. 'Escherichia Coli DNA Topoisomerase I Mutants: Increased Supercoiling Is Corrected by Mutations near Gyrase Genes'. *Cell* 31 (1): 35–42. https://doi.org/10.1016/0092-8674(82)90402-0.

Raymann, Kasie, Céline Brochier-Armanet, and Simonetta Gribaldo. 2015. 'The Two-Domain Tree of Life Is Linked to a New Root for the Archaea'. *Proceedings of the National Academy of Sciences of the United States of America* 112 (21): 6670–75. https://doi.org/10.1073/pnas.1420858112.

Raymann, Kasie, Patrick Forterre, Céline Brochier-Armanet, and Simonetta Gribaldo. 2014. 'Global Phylogenomic Analysis Disentangles the Complex Evolutionary History of DNA Replication in Archaea'. *Genome Biology and Evolution* 6 (1): 192–212. https://doi.org/10.1093/gbe/evu004.

Reyes-Domínguez, Yazmid, Gabriel Contreras-Ferrat, Jesús Ramírez-Santos, Jorge Membrillo-Hernández, and M. Carmen Gómez-Eichelmann. 2003. 'Plasmid DNA Supercoiling and Gyrase Activity in Escherichia Coli Wild-Type and RpoS Stationary-Phase Cells'. *Journal of Bacteriology* 185 (3): 1097–1100. https://doi.org/10.1128/JB.185.3.1097-1100.2003.

Rinke, Christian, Patrick Schwientek, Alexander Sczyrba, Natalia N. Ivanova, Iain J. Anderson, Jan-Fang Cheng, Aaron Darling, et al. 2013. 'Insights into the Phylogeny and Coding Potential of Microbial Dark Matter'. *Nature* 499 (7459): 431–37. https://doi.org/10.1038/nature12352.

Robinson, Philip J. J., and Daniela Rhodes. 2006. 'Structure of the "30 Nm" Chromatin Fibre: A Key Role for the Linker Histone'. *Current Opinion in Structural Biology* 16 (3): 336–43. https://doi.org/10.1016/j.sbi.2006.05.007.

Rodriguez, A. Chapin. 2002. 'Studies of a Positive Supercoiling Machine. Nucleotide Hydrolysis and a Multifunctional "Latch" in the Mechanism of Reverse Gyrase'. *The Journal of Biological Chemistry* 277 (33): 29865–73. https://doi.org/10.1074/jbc.M202853200.

Rui, Shan, and Yuk-Ching Tse-Dinh. 2003. 'Topoisomerase Function during Bacterial Responses to Environmental Challenge'. *Frontiers in Bioscience: A Journal and Virtual Library* 8 (January): d256-263. https://doi.org/10.2741/984.

Sanders, Travis J., Craig J. Marshall, and Thomas J. Santangelo. 2019. 'The Role of Archaeal Chromatin in Transcription'. *Journal of Molecular Biology*, May. https://doi.org/10.1016/j.jmb.2019.05.006.

Sandman, K., J. A. Krzycki, B. Dobrinski, R. Lurz, and J. N. Reeve. 1990. 'HMf, a DNA-Binding Protein Isolated from the Hyperthermophilic Archaeon Methanothermus Fervidus, Is Most Closely Related to Histones'. *Proceedings of the National Academy of Sciences of the United States of America* 87 (15): 5788–91. https://doi.org/10.1073/pnas.87.15.5788.

Schoeffler, Allyn J., and James M. Berger. 2008. 'DNA Topoisomerases: Harnessing and Constraining Energy to Govern Chromosome Topology'. *Quarterly Reviews of Biophysics* 41 (1): 41–101. https://doi.org/10.1017/S003358350800468X.

Schones, Dustin E., Kairong Cui, Suresh Cuddapah, Tae-Young Roh, Artem Barski, Zhibin Wang, Gang Wei, and Keji Zhao. 2008. 'Dynamic Regulation of Nucleosome Positioning in the Human Genome'. *Cell* 132 (5): 887–98. https://doi.org/10.1016/j.cell.2008.02.022.

Sehgal, S. N., M. Kates, and N. E. Gibbons. 1962. 'Lipids of Halobacterium Cutirubrum'. *Canadian Journal of Biochemistry and Physiology* 40 (January): 69–81.

Seitz, Kiley W., Nina Dombrowski, Laura Eme, Anja Spang, Jonathan Lombard, Jessica R. Sieber, Andreas P. Teske, Thijs J. G. Ettema, and Brett J. Baker. 2019. 'Asgard Archaea Capable of Anaerobic Hydrocarbon Cycling'. *Nature Communications* 10 (1): 1822. https://doi.org/10.1038/s41467-019-09364-x.

Seitz, Kiley W., Cassandre S. Lazar, Kai-Uwe Hinrichs, Andreas P. Teske, and Brett J. Baker. 2016. 'Genomic Reconstruction of a Novel, Deeply Branched Sediment Archaeal Phylum with Pathways for Acetogenesis and Sulfur Reduction'. *The ISME Journal* 10 (7): 1696–1705. https://doi.org/10.1038/ismej.2015.233.

Shogren-Knaak, Michael, Haruhiko Ishii, Jian-Min Sun, Michael J. Pazin, James R. Davie, and Craig L. Peterson. 2006. 'Histone H4-K16 Acetylation Controls Chromatin Structure and Protein Interactions'. *Science (New York, N.Y.)* 311 (5762): 844–47. https://doi.org/10.1126/science.1124000.

Siliakus, Melvin F., John van der Oost, and Servé W. M. Kengen. 2017. 'Adaptations of Archaeal and Bacterial Membranes to Variations in Temperature, PH and Pressure'. *Extremophiles: Life Under Extreme Conditions* 21 (4): 651–70. https://doi.org/10.1007/s00792-017-0939-x.

Sinden, R. R., J. O. Carlson, and D. E. Pettijohn. 1980. 'Torsional Tension in the DNA Double Helix Measured with Trimethylpsoralen in Living E. Coli Cells: Analogous Measurements in Insect and Human Cells'. *Cell* 21 (3): 773–83.

Sinden, R. R., and D. E. Pettijohn. 1981. 'Chromosomes in Living Escherichia Coli Cells Are Segregated into Domains of Supercoiling'. *Proceedings of the National Academy of Sciences* 78 (1): 224–28. https://doi.org/10.1073/pnas.78.1.224.

Soliman, Gaber S. H., and Hans G. Trüper. 1982. 'Halobacterium Pharaonis Sp. Nov., a New, Extremely Haloalkaliphilic Archaebacterium with Low Magnesium Requirement'. *Zentralblatt Für Bakteriologie Mikrobiologie Und Hygiene: I. Abt. Originale C: Allgemeine, Angewandte Und Ökologische Mikrobiologie* 3 (2): 318–29. https://doi.org/10.1016/S0721-9571(82)80045-8.

Soo, Valerie W. C., and Tobias Warnecke. 2021. 'Slaying the Last Unicorn: Discovery of Histones in the Microalga Nanochlorum Eucaryotum'. *Royal Society Open Science* 8 (2): 202023. https://doi.org/10.1098/rsos.202023.

Spang, Anja, Eva F. Caceres, and Thijs J. G. Ettema. 2017. 'Genomic Exploration of the Diversity, Ecology, and Evolution of the Archaeal Domain of Life'. *Science (New York, N.Y.)* 357 (6351). https://doi.org/10.1126/science.aaf3883.

Spang, Anja, Laura Eme, Jimmy H. Saw, Eva F. Caceres, Katarzyna Zaremba-Niedzwiedzka, Jonathan Lombard, Lionel Guy, and Thijs J. G. Ettema. 2018. 'Asgard Archaea Are the Closest Prokaryotic Relatives of Eukaryotes'. *PLoS Genetics* 14 (3). https://doi.org/10.1371/journal.pgen.1007080.

Spang, Anja, Jimmy H. Saw, Steffen L. Jørgensen, Katarzyna Zaremba-Niedzwiedzka, Joran Martijn, Anders E. Lind, Roel van Eijk, Christa Schleper, Lionel Guy, and Thijs J. G. Ettema. 2015. 'Complex Archaea That Bridge the Gap between Prokaryotes and Eukaryotes'. *Nature* 521 (7551): 173–79. https://doi.org/10.1038/nature14447.

Stetter, Karl O. 2006. 'History of Discovery of the First Hyperthermophiles'. *Extremophiles: Life Under Extreme Conditions* 10 (5): 357–62. https://doi.org/10.1007/s00792-006-0012-7.

Sugino, A., N. P. Higgins, P. O. Brown, C. L. Peebles, and N. R. Cozzarelli. 1978. 'Energy Coupling in DNA Gyrase and the Mechanism of Action of Novobiocin'. *Proceedings of the National Academy of Sciences of the United States of America* 75 (10): 4838–42. https://doi.org/10.1073/pnas.75.10.4838.

Takemata, Naomichi, and Stephen D. Bell. 2020. 'Emerging Views of Genome Organization in Archaea'. *Journal of Cell Science* 133 (10). https://doi.org/10.1242/jcs.243782.

Takemata, Naomichi, and Stephen D. Bell. 2021. 'Multi-Scale Architecture of Archaeal Chromosomes'. *Molecular Cell* 81 (3): 473-487.e6. https://doi.org/10.1016/j.molcel.2020.12.001.

Takemata, Naomichi, Rachel Y. Samson, and Stephen D. Bell. 2019. 'Physical and Functional Compartmentalization of Archaeal Chromosomes'. *Cell* 179 (1): 165-179.e18. https://doi.org/10.1016/j.cell.2019.08.036.

Terui, Yusuke, Mio Ohnuma, Kaori Hiraga, Etsuko Kawashima, and Tairo Oshima. 2005. 'Stabilization of Nucleic Acids by Unusual Polyamines Produced by an Extreme Thermophile, Thermus Thermophilus'. *The Biochemical Journal* 388 (Pt 2): 427–33. https://doi.org/10.1042/BJ20041778.

Thoma, F., T. Koller, and A. Klug. 1979. 'Involvement of Histone H1 in the Organization of the Nucleosome and of the Salt-Dependent Superstructures of Chromatin'. *The Journal of Cell Biology* 83 (2 Pt 1): 403–27. https://doi.org/10.1083/jcb.83.2.403.

Tolsma, Thomas O., and Jeffrey C. Hansen. 2019. 'Post-Translational Modifications and Chromatin Dynamics'. *Essays in Biochemistry* 63 (1): 89–96. https://doi.org/10.1042/EBC20180067.

Toro Duany, Yoandris del, Stefan P. Jungblut, Andreas S. Schmidt, and Dagmar Klostermeier. 2008. 'The Reverse Gyrase Helicase-like Domain Is a Nucleotide-Dependent Switch That Is Attenuated by the Topoisomerase Domain'. *Nucleic Acids Research* 36 (18): 5882–95. https://doi.org/10.1093/nar/gkn587.

Tretter, Elsa M., Jeffrey C. Lerman, and James M. Berger. 2010. 'A Naturally Chimeric Type IIA Topoisomerase in Aquifex Aeolicus Highlights an Evolutionary Path for the Emergence of Functional Paralogs'. *Proceedings of the National Academy of Sciences of the United States of America* 107 (51): 22055–59. https://doi.org/10.1073/pnas.1012938107.

Uhlmann, Frank. 2016. 'SMC Complexes: From DNA to Chromosomes'. *Nature Reviews. Molecular Cell Biology* 17 (7): 399–412. https://doi.org/10.1038/nrm.2016.30.

Val, Marie-Eve, Martial Marbouty, Francisco de Lemos Martins, Sean P. Kennedy, Harry Kemble, Michael J. Bland, Christophe Possoz, Romain Koszul, Ole Skovgaard, and Didier Mazel. 2016. 'A Checkpoint Control Orchestrates the Replication of the Two Chromosomes of Vibrio Cholerae'. *Science Advances* 2 (4): e1501914. https://doi.org/10.1126/sciadv.1501914.

Valens, Michèle, Stéphanie Penaud, Michèle Rossignol, François Cornet, and Frédéric Boccard. 2004. 'Macrodomain Organization of the Escherichia Coli Chromosome'. *The EMBO Journal* 23 (21): 4330–41. https://doi.org/10.1038/sj.emboj.7600434.

Valenti, Anna, Giuseppe Perugino, Takehiko Nohmi, Mosè Rossi, and Maria Ciaramella. 2009. 'Inhibition of Translesion DNA Polymerase by Archaeal Reverse Gyrase'. *Nucleic Acids Research* 37 (13): 4287–95. https://doi.org/10.1093/nar/gkp386.

Vanden Broeck, Arnaud, Christophe Lotz, Julio Ortiz, and Valérie Lamour. 2019. 'Cryo-EM Structure of the Complete E. Coli DNA Gyrase Nucleoprotein Complex'. *Nature Communications* 10 (1): 4935. https://doi.org/10.1038/s41467-019-12914-y.

Vanwonterghem, Inka, Paul N. Evans, Donovan H. Parks, Paul D. Jensen, Ben J. Woodcroft, Philip Hugenholtz, and Gene W. Tyson. 2016. 'Methylotrophic Methanogenesis Discovered in the Archaeal Phylum Verstraetearchaeota'. *Nature Microbiology* 1 (October): 16170. https://doi.org/10.1038/nmicrobiol.2016.170.

Verma, Subhash C., Zhong Qian, and Sankar L. Adhya. 2019. 'Architecture of the Escherichia Coli Nucleoid'. *PLoS Genetics* 15 (12). https://doi.org/10.1371/journal.pgen.1008456.

Vogels, G. D., W. F. Hoppe, and C. K. Stumm. 1980. 'Association of Methanogenic Bacteria with Rumen Ciliates'. *Applied and Environmental Microbiology* 40 (3): 608–12.

Wagner, Alexander, Rachel J. Whitaker, David J. Krause, Jan-Hendrik Heilers, Marleen van Wolferen, Chris van der Does, and Sonja-Verena Albers. 2017. 'Mechanisms of Gene Flow in Archaea'. *Nature Reviews Microbiology* 15 (8): 492–501. https://doi.org/10.1038/nrmicro.2017.41.

Werner, Finn, and Dina Grohmann. 2011. 'Evolution of Multisubunit RNA Polymerases in the Three Domains of Life'. *Nature Reviews. Microbiology* 9 (2): 85–98. https://doi.org/10.1038/nrmicro2507.

Wierer, Sebastian, Peter Daldrop, Misbha Ud Din Ahmad, Winfried Boos, Malte Drescher, Wolfram Welte, and Ralf Seidel. 2016. 'TrmBL2 from Pyrococcus Furiosus Interacts Both with Double-Stranded and Single-Stranded DNA'. *PloS One* 11 (5): e0156098. https://doi.org/10.1371/journal.pone.0156098.

Williams, Tom A., Cymon J. Cox, Peter G. Foster, Gergely J. Szöllősi, and T. Martin Embley. 2020. 'Phylogenomics Provides Robust Support for a Two-Domains Tree of Life'. *Nature Ecology & Evolution* 4 (1): 138–47. https://doi.org/10.1038/s41559-019-1040-x.

Williams, Tom A., Gergely J. Szöllősi, Anja Spang, Peter G. Foster, Sarah E. Heaps, Bastien Boussau, Thijs J. G. Ettema, and T. Martin Embley. 2017. 'Integrative Modeling of Gene and Genome Evolution Roots the Archaeal Tree of Life'. *Proceedings of the National Academy of Sciences of the United States of America* 114 (23): E4602–11. https://doi.org/10.1073/pnas.1618463114.

Woese, C. R., and G. E. Fox. 1977. 'Phylogenetic Structure of the Prokaryotic Domain: The Primary Kingdoms'. *Proceedings of the National Academy of Sciences of the United States of America* 74 (11): 5088–90. https://doi.org/10.1073/pnas.74.11.5088.

Woese, C. R., O. Kandler, and M. L. Wheelis. 1990. 'Towards a Natural System of Organisms: Proposal for the Domains Archaea, Bacteria, and Eucarya'. *Proceedings of the National Academy of Sciences of the United States of America* 87 (12): 4576–79. https://doi.org/10.1073/pnas.87.12.4576.

Woodcock, C. L., J. P. Safer, and J. E. Stanchfield. 1976. 'Structural Repeating Units in Chromatin. I. Evidence for Their General Occurrence'. *Experimental Cell Research* 97 (January): 101–10. https://doi.org/10.1016/0014-4827(76)90659-5.

Yang, Xi, Florence Garnier, Hélène Débat, Terence R. Strick, and Marc Nadal. 2020. 'Direct Observation of Helicase-Topoisomerase Coupling within Reverse Gyrase'. *Proceedings of the National Academy of Sciences of the United States of America* 117 (20): 10856–64. https://doi.org/10.1073/pnas.1921848117.

Yutin, Natalya, and Eugene V. Koonin. 2012. 'Archaeal Origin of Tubulin'. *Biology Direct* 7 (March): 10. https://doi.org/10.1186/1745-6150-7-10.

Zaremba-Niedzwiedzka, Katarzyna, Eva F. Caceres, Jimmy H. Saw, Disa Bäckström, Lina Juzokaite, Emmelien Vancaester, Kiley W. Seitz, et al. 2017. 'Asgard Archaea Illuminate the Origin of Eukaryotic Cellular Complexity'. *Nature* 541 (7637): 353–58. https://doi.org/10.1038/nature21031.

Zechiedrich, E. L., and N. R. Cozzarelli. 1995. 'Roles of Topoisomerase IV and DNA Gyrase in DNA Unlinking during Replication in Escherichia Coli'. *Genes & Development* 9 (22): 2859–69. https://doi.org/10.1101/gad.9.22.2859.

Zeldovich, Konstantin B., Igor N. Berezovsky, and Eugene I. Shakhnovich. 2007. 'Protein and DNA Sequence Determinants of Thermophilic Adaptation'. *PLoS Computational Biology* 3 (1): e5. https://doi.org/10.1371/journal.pcbi.0030005.

Zhang, Changyi, Alex P. R. Phillips, Rebecca L. Wipfler, Gary J. Olsen, and Rachel J. Whitaker. 2018. 'The Essential Genome of the Crenarchaeal Model Sulfolobus Islandicus'. *Nature Communications* 9 (1): 4908. https://doi.org/10.1038/s41467-018-07379-4.

Zhang, ZhenFeng, Li Guo, and Li Huang. 2012. 'Archaeal Chromatin Proteins'. *Science China. Life Sciences* 55 (5): 377–85. https://doi.org/10.1007/s11427-012-4322-y.

Zillig, W., K. O. Stetter, W. Schäfer, D. Janekovic, S. Wunderl, I. Holz, and P. Palm. 1981. 'Thermoproteales: A Novel Type of Extremely Thermoacidophilic Anaerobic Archaebacteria Isolated from Icelandic Solfataras'. *Zentralblatt Für Bakteriologie Mikrobiologie Und Hygiene: I. Abt. Originale C: Allgemeine, Angewandte Und Ökologische Mikrobiologie* 2 (3): 205–27. https://doi.org/10.1016/S0721-9571(81)80001-4.

Zuckerkandl, Emile, and Linus Pauling. 1965. 'Molecules as Documents of Evolutionary History'. *Journal of Theoretical Biology* 8 (2): 357–66. https://doi.org/10.1016/0022-5193(65)90083-4.

Article I

Les topoisomérases sont des enzymes chargées de la régulation de la topologie de l'ADN dans chaque cellule vivante. Par leurs activités de clivage transitoire de l'ADN, elles maintiennent le surenroulement de l'ADN à un niveau compatible avec les transactions de l'ADN telles que la transcription et la réplication de l'ADN et assurent la décaténation des chromosomes avant la division cellulaire. Les topoisomérases sont de type I ou de type II selon qu'elles clivent un ou deux brins d'ADN. En principe, une seule topoisomérase de type I et une seule topoisomérase de type II seraient nécessaires et suffisantes pour résoudre tous les problèmes topologiques résultant de la longueur et de l'ouverture de la double-hélice d'ADN. Cependant, au sein de ces deux classes, il existe actuellement huit sous-familles de topoisomérases réparties de manière inégale entre les trois domaines de la vie, ce qui indique que l'histoire évolutive de ces enzymes est complexe. Les topoisomérases connues sont toutes des enzymes spécifiques de l'ADN et leur activité affecte fortement les processus liés à l'ADN dans les cellules. Par conséquent, leur apparition et leur évolution sont intimement liées à l'origine de l'ADN et à son évolution en tant que support de l'information génétique.

En règle générale, les topoisomérases ont la capacité de relâcher l'ADN surenroulé, mais seules deux d'entre elles peuvent également surenrouler l'ADN relâché. La reverse gyrase peut surenrouler positivement l'ADN et cette enzyme est, à ce jour, le seul marqueur connu de la thermophilie. L'ADN surenroulé positivement est plus résistant à la dénaturation et il a donc été proposé que l'activité particulière de la reverse gyrase contribue à préserver l'intégrité du génome à haute température. La gyrase, un antagoniste de la reverse gyrase, catalyse l'introduction ATP-dépendante de supertours négatifs dans l'ADN. Selon le modèle du « Twin supercoiled domain », la gyrase élimine les supertours positifs qui s'accumulent devant l'ARN polymérase en cours de transcription ou pendant la réplication de l'ADN, permettant ainsi le déroulement de ces processus essentiels. En tandem avec la topoisomérase I (Topo I),

la gyrase régule le niveau global de surenroulement dans les cellules bactériennes et même une déviation modeste de la topologie optimale de l'ADN peut être létale chez les bactéries. L'activité de la gyrase est, via le rapport ATP/ADP intracellulaire, directement liée à l'état global du métabolisme cellulaire. Dans des conditions de stress telles que le manque de nutriments, les niveaux d'ATP diminuent, la gyrase est moins active et, par conséquent, le niveau global de surenroulement de l'ADN est modifié. En réponse, l'expression de nombreux gènes (de 7 à 48 % de l'ensemble des gènes) est simultanément modifiée pour permettre une adaptation rapide à ces conditions défavorables. La capacité de la gyrase à traduire rapidement les *stimuli* environnementaux en une réponse transcriptionnelle globale appropriée est considérée comme une caractéristique clé fournissant aux bactéries la capacité de s'adapter rapidement à un environnement changeant.

Contrairement aux orthologues bactériens qui ont été largement étudiés *in vitro*, une seule gyrase archée, celle de *Thermoplasma acidophylum* a été caractérisée biochimiquement. Cette enzyme a présenté des activités *in vitro* similaires à celles des homologues bactériens, par exemple des activités de surenroulement dépendant de l'ATP et de décaténation et relaxation indépendantes de l'ATP. Les premières études ont montré que la gyrase avait une activité de surenroulement négatif *in vivo* et que cette activité était essentielle chez les méthanogènes, les halophiles et les thermoacidophiles. On ne sait pas encore si, comme chez les bactéries, l'activité gyrase, et par son intermédiaire, l'homéostasie du surerenroulement, est liée à la régulation de l'expression génétique chromosomique chez les Archées. Le gène codant pour la rhodopsine dans l'halophile extrême *Haloferax* a été fortement induit quand l'activité de gyrase a été inhibée par la novobiocine indiquant que les promoteurs sensibles au surenroulement peuvent exister dans cette archée. Chez une autre archée halophile, *Halobacterium sp.*, le traitement à la novobiocine a entraîné la dérégulation de nombreux gènes, notamment une augmentation de l'expression de la gyrase, de la

Topo VI et de la Topo I, ce qui indique l'implication de ces enzymes dans la régulation des niveaux de surenroulement de l'ADN dans cet organisme.

Des études antérieures utilisant l'échantillonnage taxonomique disponible à l'époque, ont identifié la présence de gènes codant pour la gyrase chez toutes les Euryarchées de groupe II, mais pas chez d'autres phyla d'archées connus. Il a donc été proposé que la gyrase aurait été transférée des bactéries aux archées, vraisemblablement une seule fois à la base des Euryarchées de groupe II. Ce groupe majeur au sein du superphylum des Euryarchaeota contient 7 classes distinctes d'organismes aux modes de vie variés (méthanogènes, halophiles) et, à l'exception des Archaeoglobi, ce sont toutes des archées mésophiles. L'acquisition de la gyrase par l'ancêtre thermophile des Euryarchées de groupe II a été proposée comme ayant joué un rôle clé dans l'adaptation de cette lignée à un mode de vie mésophile.

Avec la disponibilité d'un échantillonnage taxonomique plus large, il est maintenant clair que la gyrase est également présente dans des lignées d'archées distantes des Euryarchées de groupe II, comme les DPANN et les Asgard. Cela soulève la question du moment et du nombre de transferts horizontaux de gènes inter- ou intra-domaine, ainsi que la nature des adaptations associées, spécifiques ou à l'échelle du génome. La résolution de ces questions est importante pour comprendre l'évolution des archées, et aussi pour tester des hypothèses sur la régulation de l'expression des gènes par la topologie de l'ADN chez les archées, qui est encore mal comprise. Pour répondre à ces questions, nous avons échantillonné les gyrases de tous les génomes d'archées et de bactéries disponibles et avons effectué une analyse phylogénétique et comparative des séquences.

Nous avons trouvé la gyrase dans toutes les lignées des Euryarchées de groupe II, et dans plusieurs lignées des DPANN et des Asgard, élargissant ainsi les données sur la distribution de la gyrase au sein de ces archées récemment découvertes. La topologie de l'arbre des archées suggère que la gyrase a été transférée des bactéries à la base

des Euryarchées de groupe II et qu'elle s'est ensuite répandue dans d'autres lignées par le biais de transferts horizontaux secondaires de gènes. Il est intéressant de noter que nous n'avons pas pu détecter la gyrase dans le clade des TACK, ce qui suggère une incompatibilité entre la physiologie de ces organismes et l'activité de la gyrase. Fait important, nous avons constaté que l'arbre global de la gyrase présente une topologie tripartite dans laquelle les bactéries forment deux clades correspondant aux Terrabacteria et aux Gracilicutes et les Archaea sont monophylétiques. Cela suggère que la gyrase a été transférée avant la diversification des Terrabacteria et des Gracilicutes mais après la diversification des Archaea. Un tel scénario pourrait impliquer que la diversification des Archaea est antérieure à celle des Terrabacteria et des Gracilicutes. De façon remarquable, les branches séparant les bactéries des Archées sont très courtes. Cela suggère que l'acquisition de la gyrase par les archées ne s'est pas accompagnée de changements radicaux dans sa séquence et, par conséquent, que les gyrases des archées partagent des fonctions très similaires à celles des bactéries.

Article II

Avec la découverte de la structure de l'ADN, il est devenu évident que l'ouverture de la double hélice génère des contraintes mécaniques qui provoquent un surenroulement de la molécule d'ADN. Paradoxalement, de nombreuses transactions de l'ADN, tels que la transcription et la réplication, nécessitent la séparation des brins et conduisent naturellement à un surenroulement et à un enchevêtrement de l'ADN. Ces contraintes topologiques finissent par empêcher le déroulement de ces processus cellulaires essentiels et, si elles ne sont pas résolues, conduisent à la mort cellulaire. Pour faire face à ce problème, toutes les cellules s'appuient sur les topoisomérases, une classe d'enzymes ubiquitaire qui introduisent des cassures transitoires dans l’ADN pour résoudre les intermédiaires topologiques défavorables sans endommager le génome. Les topoisomérases sont dites de type I ou type II, selon qu'elles clivent respectivement un ou deux brins d'ADN. De multiples sous-classes de chaque type de topoisomérase,

existent dans la biosphère et cette diversité a rendu particulièrement difficile l'étude de l'histoire évolutive des topoisomérases. Une énigme de longue date est de comprendre pourquoi tant de topoisomérases sont apparues au cours de l'évolution et quel rôle elles ont joué dans l'évolution des cellules.

L'ADN gyrase (ci-après gyrase), une topoisomérase de type II A, est la seule enzyme connue capable de surenrouler négativement l'ADN en utilisant l'énergie de l'hydrolyse de l'ATP. Cible importante des antibiotiques, la gyrase est essentielle et ubiquitaire chez les bactéries où elle contrôle (avec la Topo I) la densité de surenroulement des chromosomes en introduisant des supertours négatifs dans l'ADN et en relâchant les supertours positifs qui s'accumulent devant les ADN et ARN polymérases en mouvement. La contribution de la gyrase au maintien du chromosome dans un état de surenroulement négatif dans les cellules bactériennes peut avoir un impact profond sur la liaison des protéines régulatrices, la dynamique d'activation des promoteurs, la réplication de l'ADN et l'architecture des chromosomes.

Le surenroulement de l'ADN est utilisé dans un large éventail de bactéries pour transmettre rapidement les signaux environnementaux vers le chromosome et ce processus est conservé dans des espèces bactériennes phylogénétiquement éloignées. La voie la plus clairement décrite implique la modulation de l'activité de la gyrase en réponse au rapport [ATP]/[ADP] dans la cellule. Lorsque ce rapport est faible, l'activité de surenroulement de la gyrase est significativement réduite et le niveau d'expression de nombreux gènes est simultanément modifié. L'inhibition de l'activité de surenroulement de la gyrase par les antibiotiques quinolones a un effet similaire ; l'expression de jusqu'à 48% des gènes peut être dérégulée. Ce mécanisme relativement simple, rapide et général a été suggéré comme étant l'une des inventions évolutives clé permettant aux bactéries d'occuper une grande variété d'environnements.

Chez les archées, la gyrase est présente chez tous les membres d'un groupe monophylétique très diversifié (nommé cluster II Euryarchaeota par Adam et collègues)

contenant sept phyla distincts aux modes de vie très différents (acidophiles, halophiles, méthanogènes entre autres) et, sporadiquement, chez les superphyla DPANN et Asgard. Les analyses phylogénétiques initiales ont indiqué que les gyrases d'archées sont d'origine bactérienne et ont été acquises par une archée hyperthermophile grâce à un transfert horizontal de gènes. Des analyses ultérieures incluant d'autres lignées d'archées ont suggéré que ce transfert n'a eu lieu qu'une seule fois à la base des Euryarchées de groupe II. Puisque le surenroulement négatif facilite la dénaturation de l'ADN, il a été proposé que l'acquisition de la gyrase a un impact profond sur tous les processus dépendant de l'ADN, avec des conséquences importantes pour l'évolution des archées réceptrices. Cependant, comment et pourquoi la gyrase s'est fixée dans les lignées d'archées reste obscur.

L'établissement réussi de la gyrase bactérienne dans les archées est particulièrement intrigant, car l'introduction d'une activité de surenroulement négatif incontrôlée pourrait potentiellement interférer avec tous les processus liés à l'ADN. Pour aggraver les choses, l'archée réceptrice était probablement thermophile et codait donc une reverse gyrase. Cette topoisomérase a une activité de surenroulement positif, opposée à celle de la gyrase, et elle est essentielle pour la vie à haute température. Enfin, les archées possèdent la Topo VI comme principale topoisomérase de type II et son rôle prédit *in vivo* dans la relaxation des surenroulements positifs est redondant avec celui de la gyrase.

Pour mieux comprendre les défis imposés par la gyrase à une cellule archéenne naïve, nous avons introduit une gyrase bactérienne dans l'archée hyperthermophile *Thermococcus kodakarensis* TS559, génétiquement modifiable, qui possède naturellement un ADN surenroulé positivement. Cette archée possède des histones et trois topoisomérases (reverse gyrase, Topo III et Topo VI), imitant ainsi le "kit" topologique présent chez l'ancêtre des Euryarchaeota de groupe II. Comme donneur de gyrase, nous avons choisi celle de la bactérie *Thermotoga maritima* (TmGyrAB) puisque les orthologues les plus proches sont des gyrases d'archées, ce qui augmente

les chances que son activité ne soit pas altérée par le contexte cellulaire archée. De plus, les températures optimales de croissance de *T. maritima* (80°C) et *T. kodakarensis* (85°C) sont similaires, les deux ayant été isolées de sources hydrothermales marines et elles peuvent être co-cultivées en laboratoire. Enfin, TmGyrAB présente l'activité de surenroulement négatif attendue à la fois *in vitro* et *in vivo*.

Nous avons montré que la gyrase de *T. maritima* était active : l'ADN plasmidique normalement surenroulé positivement est converti en ADN fortement surenroulé négativement dans *T. kodakarensis*. La gyrase interagit avec le génome de *T. kodakarensis* et altère l'expression de centaines de gènes, dont une partie répond spécifiquement à l'activité de surenroulement négatif. La reverse gyrase est la seule topoisomérase surexprimée, bien que modestement, en réponse à l'activité de surenroulement négatif. Malgré les changements topologiques non naturels induits par la gyrase, la croissance de *T. kodakarensis* n'est pas affectée. Nous concluons que la gyrase est remarquablement bien tolérée par *T. kodakarensis*, ce qui suggère l'existence de mécanismes de résilience contre le stress topologique qui ont pu être déterminants pour l'établissement de la gyrase dans le domaine des Archées au cours de l'évolution.

Article III

Les hautes températures représentent un défi considérable pour la physiologie cellulaire et en particulier pour la stabilité du génome. Comme le mouvement moléculaire augmente proportionnellement à l'élévation de la température, les molécules, les membranes et l'ADN deviennent instables. Concrètement, dans des conditions *in vitro*, l'exposition de l'ADN à des températures élevées entraîne des taux élevés d'oxydation, d'alkylation, de désamination, de dépurination et enfin de cassures de brins. Pour protéger leurs génomes contre ces lésions toxiques de l'ADN, les thermophiles et les hyperthermophiles ont recours à des adaptations spécifiques. Entre autres, ils hébergent des ARNr riches en GC, des lipides spécifiques, des proportions

élevées d'acides aminés thermostables dans leurs protéines (typiquement I, V, Y, W R, E, L), des protéines chaperones, des concentrations élevées de sels et de polyamines dans leurs cytoplasmes, des transporteurs d'ions spécifiques et une topoisomérase, la reverse gyrase.

Ce type particulier de topoisomérase IA a été découvert en 1984 chez *Sulfolobus acidocaldarius* et nommé après son activité unique de surenroulement positif, opposé à celle de la gyrase. La reverse gyrase est une fusion des domaines SF2-hélicase et topoisomérase, et présente donc une activité ATP-dépendante et $MgCl_2$-dépendante.

Il a été proposé que la reverse gyrase crée un surenroulement négatif et positif en ouvrant l'ADN, puis en relâchant uniquement le surenroulement négatif, ce qui entraîne une augmentation nette du surenroulement positif. Les domaines hélicase et topoisomérase coopèrent grâce au « latch » au cours du processus. De façon remarquable, jusqu'à présent, la reverse gyrase est la seule protéine présente chez tous les hyperthermophiles et la plupart des thermophiles, ce qui en fait la marque de l'hyperthermophilie. On pensait que son effet protecteur contre les hautes températures provenait de la plus grande stabilité de l'ADN surenroulé positivement, une topologie qui présente un empilement accru des bases et qui est plus résistante à la dénaturation des deux brins à haute température.

La compaction de l'ADN par la chromatinisation est une autre stratégie de stabilisation de l'ADN à haute température. Les histones des archées sont classiquement organisées en nucléosomes tétramériques. Mais certaines archées comme *T. kodakarensis*, possèdent des histones qui peuvent étendre cette structure de base par plusieurs dimères d'histones. Il en résulte un assemblage d'histones en oligomères de différentes tailles qui enrobent l'ADN dans une structure filamenteuse. Il a été démontré que la couverture en histone du génome de *T. kodakarensis* est presque complète. Cette couverture élevée en histones suggère que les histones sont des acteurs majeurs de l'architecture du génome et de la régulation des gènes chez *T. kodakarensis*. De

manière importante, les histones peuvent réguler le surenroulement à la fois en le contraignant ou en libérant des supertours de sens opposé. Inversement, la liaison des histones à l'ADN et l'oligomérisation des histones peuvent être affectées par le surenroulement de l'ADN.

Chez la plupart des hyperthermophiles étudiés jusqu'à présent, les plasmides ont été trouvés dans un état de surenroulement relâché ou légèrement positif. Cependant, les hyperthermophiles contenant de l'ADN négativement surenroulé existent et leurs génomes codent invariablement pour la gyrase, indiquant que lorsque les deux topoisomérases coexistent dans une cellule, la gyrase dicte la topologie globale de l'ADN. Nous avons récemment démontré, en introduisant une gyrase active dans l'archée hyperthermophile naïve *Thermococcus kodakarensis*, que la topologie des plasmides peut être convertie d'un surenroulement positif à un surenroulement négatif sans effet négatif sur la croissance de cette archée. Ce résultat remet en question l'hypothèse selon laquelle le surenroulement positif de l'ADN est une caractéristique importante des hyperthermophiles et soulève des questions sur le rôle physiologique de la reverse gyrase.

L'étude de la fonction de la reverse gyrase chez les archées hyperthermophiles par des approches génétiques a produit des résultats ambigus. La délétion du gène codant pour la reverse gyrase chez l'euryarchée *T. kodakarensis* est viable et ne provoque pas de défaut de croissance à 85°C, sa température de croissance optimale. Cependant, à 90°C, le taux de croissance est divisé par deux et la souche ne croît plus du tout à 93°C. Chez *Pyrococcus furiosus*, une espèce étroitement apparentée, l'absence de reverse gyrase est létale à 95°C, une température de croissance sous-optimale pour cette archée qui croît de manière optimale à 100°C. Chez l'archée *Sulfolobus islandicus*, chacun des deux gènes codant pour la reverse gyrase peut être supprimé séparément, mais la suppression de TopR2 ne peut être obtenue qu'après 14 à 20 jours d'incubation ce qui semble difficilement compatible avec la survie dans des conditions de croissance naturelles.

Outre son activité de surenroulement positif, de plus en plus d'études suggèrent l'implication directe de la reverse gyrase dans la réparation de l'ADN. La reverse gyrase de *Sulfolobus solfataricus* forme des complexes covalents stables avec l'ADN endommagé *in vitro* par irradiation aux UV, en utilisant l'enzyme purifiée ou des extraits cellulaires. À un rapport protéine/ADN élevé, la reverse gyrase d'*Archaeoglobus fulgidus* empêche l'agrégation inappropriée de régions d'ADN dénaturées par la chaleur et favorise un appariement correct *in vitro*, protégeant ainsi l'ADN des cassures double brin induites par la chaleur, indépendamment de toute activité enzymatique. La reverse gyrase interagit avec la protéine de liaison à l'ADN simple brin RPA de *Pyrococcus abyssi*. Chez *Sulfolobus solfataricus*, il a été démontré que la reverse gyrase forme un complexe avec la protéine de liaison à l'ADN simple brin SSB et l'ADN. La liaison à la SSB augmente l'activité topoisomérase de la reverse gyrase, ce qui suggère son implication directe dans la réparation de l'ADN. Dans une autre étude, le même groupe a montré que la reverse gyrase co-immunoprécipite avec la polymérase translésionnelle Y (PolY) connue pour son rôle essentiel dans la réparation de l'ADN. La reverse gyrase inhibe l'activité de PolY *in vitro* et l'activité de surenroulement positif est requise pour l'inhibition, ce qui suggère que la reverse gyrase peut réguler l'activité de la PolY lorsque l'ADN est endommagé. D'autres expériences ont étudié le rôle de la reverse gyrase lorsque des cellules de *Sulfolobus islandicus* sont exposées au méthyl-méthane-sulfonate (MMS), un agent alkylant de l'ADN qui provoque des cassures de l'ADN. La déplétion de TopR1 dans ces conditions a entraîné une dégradation accélérée de l'ADN génomique accompagnée d'un taux plus élevé de mort cellulaire.

Depuis qu'il a été montré il y a 20 ans que la reverse gyrase est la seule protéine caractéristique de l'hyperthermophilie, un travail considérable a été réalisé pour comprendre ce qui la rend indispensable à la vie à haute température. En dépit de ces efforts, le rôle physiologique de la reverse gyrase et, en particulier, sa contribution précise au maintien de l'intégrité du génome restent obscurs. L'une des principales questions non résolues reste de comprendre quelle est la contribution de la reverse

gyrase au surenroulement global de l'ADN chromosomique. Pour répondre à cette question, nous avons caractérisé *in vivo* l'impact de la délétion de la reverse gyrase sur la topologie de l'ADN cellulaire. Nous avons d'abord démontré que la reverse gyrase est effectivement responsable du surenroulement positif des plasmides chez l'archée hyperthermophile *Thermococcus kodakarensis*. En utilisant la digestion de la chromatine par la nucléase micrococale (MNase), nous avons montré que la délétion de la reverse gyrase n'entraîne pas de profil aberrant, ce qui suggère que l'absence d'activité de surenroulement positif dans les cellules de *T. kodakarensis* n'a pas d'effet important sur l'assemblage des nucléosomes.

Article IV

La nature hélicoïdale de l'ADN, tout en assurant la stabilité de la molécule, impose de fortes contraintes topologiques chaque fois que les deux brins doivent être séparés. Cela signifie que les transactions de l'ADN comme la réplication, la transcription, la recombinaison et la réparation complexifient la topologie de l'ADN, entraînant la formation de supertours, de nœuds, de caténanes et de formes non-B de l'ADN. L'accumulation de ces structures influence en retour les transactions de l'ADN, interférant avec les processus cellulaires essentiels. Pour maintenir les contraintes topologiques à un niveau physiologique, les cellules s'appuient sur un large éventail d'acteurs dédiés à la topologie de l'ADN qui coopèrent pour relâcher les contraintes ou les contenir dans des régions spécifiques des chromosomes. La structure topologique de l'ADN la plus courante est le surenroulement, qui est directement régulé par les topoisomérases.

Le surenroulement de l'ADN a été largement étudié chez les organismes modèles en étudiant la topologie des plasmides à l'aide d'électrophorèse en gradient de sucrose ou en gel d'agarose. Ces méthodes sont basées sur la mobilité accrue des plasmides surenroulés (par rapport aux plasmides relâchés) du fait de leur compaction. Bien que faciles à mettre en œuvre, ces méthodes sont limitées par le fait que les plasmides, en

raison de leur petite taille, ne représentent pas fidèlement les chromosomes beaucoup plus grands et complexes. Des approches génétiques, reposant sur la recombinaison spécifique des sites de la resolvase $\gamma\delta$ ou de l'intégrase du bactériophage λ, ont été utilisées dans le passé pour estimer la distribution du surenroulement à travers le chromosome. Cependant, ces méthodes sont limitées aux bactéries modèles génétiquement modifiables et impliquent de potentiels effets polaires. D'autres méthodes basées sur des dérivés du psoralène permettent de surmonter ces limites en mesurant le surenroulement relatif directement sur les chromosomes sans avoir recours à la génétique. Le psoralène est une petite molécule aromatique plane qui s'intercale entre les paires de bases de l'ADN et établit des ponts covalents intramoléculaires sous l'effet d'une exposition aux UV. L'ADN surenroulé négativement incorpore environ deux fois plus de psoralène que l'ADN relâché, ce qui en fait un outil fiable pour cartographier le surenroulement négatif de l'ADN. Au cours de la dernière décennie, le psoralène a été utilisé pour étudier le surenroulement à l'échelle du génome chez divers modèles de bactéries et d'eucaryotes mais pas chez les Archées. Par conséquent, l'état de surenroulement des chromosomes n'est pas connu chez les Archées.

Le surenroulement négatif de l'ADN a un rôle crucial dans la régulation des transactions de l'ADN par sa capacité à favoriser l'ouverture de la double hélice. D'autre part, le surenroulement positif de l'ADN stabilise la double hélice dans une conformation fermée, ce qui diminue son potentiel de dénaturation. Chez les bactéries, le surenroulement est utilisé comme un régulateur transcriptionnel rapide et massif par sa capacité à changer simultanément l'expression d'un grand nombre de gènes en réponse à des changements environnementaux. Cette régulation est principalement basée sur l'activité de la gyrase qui introduit un surenroulement négatif dans le génome proportionnellement au rapport [ATP]/[ADP] dans la cellule. Par conséquent, les génomes bactériens sont en moyenne surenroulés négativement mais la distribution du surenroulement négatif n'est pas homogène et varie en fonction de

l'origine de réplication et de la phase de croissance. Chez *Escherichia coli*, la seule bactérie pour laquelle le surenroulement a été cartographié à l'échelle du génome, un gradient décroissant de surenroulement négatif est mesuré du terminus vers l'origine de réplication en phase stationnaire alors qu'aucun gradient n'a été observé lorsque la bactérie croît de manière exponentielle. Il a été démontré que ce schéma spécifique est corrélé à l'intensité de la transcription et aux sites de liaison de la gyrase, ce qui reflète l'importance de cette topoisomérase pour une transcription efficace et un contrôle global du surenroulement. Les eucaryotes sont dépourvus de gyrase, mais possèdent également un génome surenroulé négativement. Ce surenroulement négatif est généré par la fixation des histones sur l'ADN. Le contrôle de l'expression génétique à partir de ces *loci* contraints est généré par une relation complexe entre la transcription, la liaison des histones et la relaxation de l'ADN par les topoisomérases. Il est important de noter que chez les eucaryotes, les supertours négatifs libres sont concentrés près des sites de début de transcription (TSS) et des sites de fin de transcription (TTS) des gènes actifs, tandis que les supertours positifs libres sont plutôt situés au milieu des gènes activement transcrits. Une telle organisation favorise les structures d'ADN non-B aux limites des gènes ce qui les isolent du reste du génome, augmentant le recyclage de l'ARN polymérase du TSS au TTS et maintenant les corps des gènes en surenroulement positif pour éviter le repositionnement des nucléosomes.

Les études de la topologie globale de l'ADN chez les archées n'en sont qu'à leurs débuts, mais les premières études de la topologie des plasmides *in vivo* indiquent que, comme c'est souvent le cas, ces organismes présentent des caractéristiques idiosyncrasiques que l'on ne retrouve pas chez les bactéries ou les eucaryotes. Par exemple, l'ADN plasmidique positivement surenroulé n'est jusqu'à présent trouvé que dans les lignées d'archées thermophiles. Cette spécificité est censée protéger l'intégrité du génome à haute température, mais elle affecte probablement aussi la géométrie globale de l'ADN et la régulation de l'expression des gènes. Au-delà du cas particulier des thermophiles, les Archées possèdent des génomes petits et denses avec la majorité

des gènes structurés en opérons et leur transcription et traduction sont couplées. Ces propriétés bactériennes sont contrebalancées par une machinerie de transcription homologue à celle des eucaryotes et la présence d'histones dans la plupart des lignées d'archées. Il est probable que l'intégration réussie au cours de l'évolution d'une telle combinaison mosaïque de caractéristiques moléculaires ait abouti à l'émergence de mécanismes originaux de contrôle de la topologie de l'ADN chez les archées.

Dans cette étude, nous avons utilisé le pontage de l'ADN par le triméthylpsoralène (TMP) couplée à un séquençage de nouvelle génération (NGS) pour établir la première carte du surenroulement de l'ADN chez une archée. Nous nous sommes concentrés sur *Thermococcus kodakarensis*, une archée hyperthermophile génétiquement modifiable avec un kit topologique relativement simple composé de deux histones (HtkA et HtkB), deux NAPs (TrmBL2 et Alba), une SMC et trois topoisomérases (Topo III, Topo VI et reverse gyrase). Nous avons utilisé quatre souches construites précédemment : une souche TKAg de référence, une souche témoin TKY119F exprimant une gyrase inactive, une souche exprimant la gyrase TKgyrAB et un mutant de délétion de la reverse gyrase TKΔRG. En comparant l'enrichissement relatif des pontages au TMP sur le génome de ces souches, nous prévoyons de (i) cartographier la distribution du surenroulement négatif dans *T. kodakarensis*, donnant un aperçu de l'interaction entre la topologie locale de l'ADN, la transcription des gènes et l'architecture du chromosome ; (ii) établir l'importance de la reverse gyrase pour le surenroulement global du chromosome ; (iii) décrire plus en détail l'impact de l'acquisition de la gyrase sur la topologie du chromosome d'un organisme naïf.

université PARIS-SACLAY
ÉCOLE DOCTORALE
Structure et dynamique des systèmes vivants (SDSV)

Titre : Étude *in vivo* de la topologie de l'ADN chez l'archée hyperthermophile *Thermococcus kodakarensis*

Mots clés : ADN gyrase, reverse gyrase, topologie de l'ADN, *Thermococcus kodakarensis*, pontage au psoralène

Résumé : Au sein des cellules, l'ADN existe dans un état surenroulé. Les topoisomérases régulent le surenroulement de l'ADN pour permettre aux processus essentiels, comme la transcription ou la réplication, d'avoir lieu. Au cours de ma thèse je me suis intéressé à la gyrase qui est unique par sa capacité de surenrouler l'ADN négativement. Des études antérieures suggèrent qu'un transfert horizontal de gènes depuis les bactéries a permis l'implantation de la gyrase chez les Archées. L'acquisition d'une telle activité doit avoir profondément bouleversé la topologie de l'ADN d'archée receveuse. Comment cette archée « naïve » a pu intégrer une gyrase et quels ont été les avantages évolutifs associés ? Pour aborder cette question, j'ai mis en œuvre une approche intégrative, mêlant analyses phylogénétiques, génie génétique et techniques d'analyse du génome dans sa globalité. J'ai montré que la gyrase a été acquise tardivement dans l'histoire des Archées et que cet unique événement a été suivi par une expansion limitée *via* de rares transferts secondaires. Pour mimer le transfert ancestral, j'ai exprimé une gyrase bactérienne dans l'archée *Thermococcus kodakarensis*, qui en est dépourvue et présente un surenroulement de l'ADN positif. J'ai pu démontrer que la gyrase introduit un surenroulement négatif important dans l'ADN de *T. kodakarensis*. Ce bouleversement topologique dérégule l'expression de nombreux gènes, mais pas les topoisomérases endogènes, et n'a pas d'impact sur la croissance de *T. kodakarensis*. Cela indique l'existence chez cette archée, de mécanismes permettant de résister efficacement au stress topologique. Pour tester cette hypothèse, j'ai mis au point la technique de pontage au psoralène, qui permet de cartographier le surenroulement directement sur le chromosome, chez *T. kodakarensis*. Couplées à un séquençage à haut débit, ces données permettront pour la première fois de décrire en détail la répartition du surenroulement à l'échelle du chromosome chez une archée.

Title : *In vivo* study of DNA topology in hyperthermophilic archaeon *Thermococcus kodakarensis*

Keywords : DNA gyrase, reverse gyrase, DNA topology, *Thermococcus kodakarensis*, psoralen crosslink

Abstract : In all cells the DNA exists in supercoiled form. Topoisomerases regulate the DNA supercoiling allowing essential processes such as transcription or replication to occur. In course of my thesis I focused on DNA gyrase which is unique by its capacity to supercoil DNA negatively. Previous studies suggested that this enzyme was transferred from Bacteria to Archaea via a single HGT. This acquisition must have had completely changed the DNA topology of the recipient archaeon, thus impacting all DNA transactions. I therefore asked how this naïve archaeon could accommodate an active gyrase and what was the associated evolutive advantage? To tackle this question I undertook an integrative approach using phylogenetic analysis, genetic engineering and genome wide analyses.
I showed that gyrase was acquired relatively late in the history of Archaea and that this event was followed by a limited expansion through rare secondary transfers. To mimic the ancestral transfer, I expressed a bacterial gyrase in the naïve archaeon *Thermococcus kodakarensis*, which naturally has positively supercoiled DNA. I demonstrated that the gyrase was active in *T. kodakarensis* leading to strong negative supercoiling of plasmid DNA. Surprisingly, the drastic change of DNA topology, while mildly deregulating the expression of hundreds of genes, did not affect growth of *T. kodakarensis*. Notably, the endogenous topoisomerases were not deregulated suggesting existence of uncommon resilience mechanisms against torsional stress in this archaeon. To test this hypothesis, I established psoralen crosslinking assay to directly map supercoiling on the chromosome of *T. kodakarensis*. Coupled with next generation sequencing, this technique will allow us to map for the first time the distribution of supercoiling at the chromosome scale in an archaeon. My thesis work now provides the necessary tools to investigate how DNA supercoiling is distributed and maintained in hyperthermophiles and in archaea in general.

Maison du doctorat de l'Université Paris-Saclay
2ème étage aile ouest, Ecole normale supérieure Paris-Saclay
4 avenue des Sciences,
91190 Gif sur Yvette, France