Mechanism of ColE1 plasmid replication regulation by the Hfq protein of Escherichia coli

Krzysztof Kubiak

DNA replication is one of the basic cellular processes necessary for the reproduction or multiplication of the vast majority of organisms (with the exceptions of viroids and viruses with RNA genomes). Consequently, identifying the mechanisms that regulate the control of this process is crucial to understanding the basic biological rules that ensure the function, survival, and reproduction of both prokaryotic and eukaryotic cells. In bacterial cells, nucleoid replication is essential for survival, as any factors that significantly interfere with this process ultimately cause their death.

Escherichia coli, a Gram-negative bacterium, is one of the best-known organisms studied to date. Its chromosome replication is controlled by a complex network of regulators to ensure the correct inheritance of genetic information. In addition, a number of extrachromosomal genetic elements, such as plasmids and bacteriophages, which are capable of autonomous duplication in *E. coli* cells, serve as models for studying the mechanisms of essential cellular processes such as DNA replication. Therefore, on the one hand, it is possible to conduct advanced research at the molecular level with the use of prokaryotic models, and on the other hand, the obtained results of such research can usually be regarded as results of general biological significance. In addition, it should be remembered that many plasmids carry antibiotic resistance genes, hence they are of great importance in medicine, and learning about their functioning may help in the search for new antibacterial drugs. Moreover, plasmids play a huge role in biotechnology as they are often used as expression vectors in the construction of strains that efficiently produce recombinant proteins.

The *E. coli* Hfq protein was discovered as an essential factor in the development of the bacteriophage Qβ, whose genome is made up of RNA. The *hfq* gene orthologs are present in almost half of the so far sequenced bacterial genomes and in some archaeal species, while structural and functional homologues are present in most eukaryotes. This suggests an important and evolutionarily conserved role for the product of this gene. In *E. coli* cells, most Hfq protein molecules are present in the cytoplasmic fraction,

mainly associated with the ribosomes, but some molecules interact with the cell membrane and nucleoid. Hfq nucleoid fraction is as large as 10–20% which represents about 5% of the total nucleoid-associated proteins, while its cytoplasmic and membrane-bound fractions are about 30 and 50%, respectively. The Hfq protein has been described as multifunctional. Among other things, it regulates RNA metabolism and interacts with proteins related to RNA degradation, e.g. poly(A) polymerase I (PAP I), PNPase and RNase E. The participation of the Hfq protein in the processes related to the virulence of pathogenic bacteria has also been described. Interesting is the observed function of Hfq in nucleoid organization and its interactions with DNA. The deletion of the *hfq* gene manifests in various changes in cell physiology, which proves the pleiotropic effect of this protein.

The results of the preliminary research of the teams led by the supervisors of my doctoral dissertation, Prof. Grzegorz Węgrzyn and Prof. Veronique Arluison, suggested the intriguing possibility that Hfq may be involved - directly or indirectly - in the regulation of DNA replication. Initial experiments were carried out with the use of various plasmids, and the most interesting effects were observed with the replicons belonging to the ColE1 group. Hence, the main aim of this dissertation was to understand the mechanisms of regulation of the replication of ColE1-like plasmids by the Hfq protein. Replication of these plasmids begins in the *origin* region, and two transcripts, called RNA I and RNA II, are involved in regulating this process. The pre-primer RNA II transcript creates an RNA-DNA hybrid in the *origin* region, and after its RNA strand is cut by RNase H, a primer is formed, which can be extended by DNA polymerase I, which is the beginning of the ColE1 plasmid replication process. The negative regulator of replication is RNA I, an RNA molecule partially complementary to RNA II, forming a hybrid with it, which prevents RNA II from hybridizing with DNA. In addition, the interactions of RNA I with RNA II are enhanced by the Rom protein encoded by the plasmid. At the time of starting my doctoral dissertation, the role of the Hfq protein in this process was unknown, and the only experiments indicating that this protein could be involved in the control of ColE1-like plasmid replication were based on the analysis of the effectiveness of *hfq* mutant transformation using such a plasmid.

The first article included in this dissertation (**Front. Mol. Biosci. 2016; 3: 36**) is a review, describing the above-mentioned properties of the Hfq protein in more detail. In particular, attention was paid to the interactions between Hfq and DNA molecules, as well as

the possibility of participation of this protein in various processes related to DNA functions, including plasmid replication. This work is therefore a theoretical introduction to the entire series of articles.

The second review in the series of publications (**Methods Mol. Biol. 2022; in press**) is a methodological article that describes the techniques used to test involvement of bacterial amyloid proteins in the antibiotic resistance processes. The methods described in this work were presented on models of the Hfq protein, the C-terminal domain of which forms amyloid structures, as well as of plasmids carrying antibiotic resistance genes. In this light, it is an introduction to *in vivo* research, which are then included in subsequent publications that make up the series of publications presented as my doctoral dissertation.

I started my research from determining the possibility of interaction of the Hfq protein with DNA and the effects of such interactions on the structure of this nucleic acid. In the article describing this research (**Nucleic Acids Res. 2017; 45: 7299-7308**), various experimental methodologies, including fluorescence microscopy imaging of single DNA molecules confined inside nanofluidic channels, atomic force microscopy, isothermal titration microcalorimetry and electrophoretic mobility assays have been used to follow the assembly of the C-terminal and N-terminal regions of Hfq on DNA. Results highlighted the role of Hfq's C-terminal domain in DNA binding, change in mechanical properties of the double helix and compaction of DNA into a condensed form. The propensity for bridging and compaction of DNA by the C-terminal domain might be related to aggregation of bound protein and may have implications for protein binding related gene regulation and DNA replication control.

Another experimental work (**Int. J. Mol. Sci. 2021; 22: 8886**) concerned the role of the Hfq protein and its C-terminal domain in the regulation of bacterial resistance to antibiotics, related to the presence of appropriate genes on plasmids, including ColE1-type plasmids. It was found that the presence of Hfq is required for survival of plasmid-containing *Escherichia coli* cells against high concentrations of chloramphenicol, tetracycline and ampicillin, as *hfq*$^{+}$ bacteria were more resistant to these antibiotics than the *hfq*-null mutant. In striking contrast, production of Hfq resulted in low resistance to high concentrations of kanamycin when the antibiotic-resistance marker was chromosome-borne, with deletion of *hfq* resulting in increasing bacterial survival. These results were observed both in solid and

liquid media, suggesting that antibiotic resistance is an intrinsic feature of these strains rather than a consequence of adaptation. Despite its major role as RNA chaperone, which also affects mRNA stability, Hfq was not found to significantly affect *kan* and *tet* mRNAs turnover. Nevertheless, *kan* mRNA steady-state levels were higher in the *hfq*-null mutant compared to the *hfq*$^{+}$ strain, suggesting that Hfq can act as a repressor of *kan* expression. This observation does correlate with the enhanced resistance to high levels of kanamycin observed in the *hfq*-null mutant. Furthermore, dependency on Hfq for resistance to high doses of tetracycline was found to depend on plasmid copy number, which was only observed when the resistance marker was expressed from a low copy plasmid (pSC101) but not from a medium copy plasmid (ColE1-like plasmid pBR322). This suggests that Hfq may influence survival against high doses of antibiotics through mechanisms that remain to be determined. I confirmed that efficiency of transformation of *hfq* mutants with ColE1-like plasmids is decreased relative to wild-type bacteria. Studies with pBR322Δrom may also suggest an interplay between Hfq and Rom in the regulation of ColE1-like plasmid replication. Results of experiments with a mutant devoid of the part of the *hfq* gene coding for the C-terminal region of Hfq suggested that this region, as well as the N-terminal region, may be involved in the regulation of expression of antibiotic resistance in *E. coli* independently.

Synchrotron Radiation Circular Dichroism (SRCD) was used to thoroughly investigate the interactions of the C-terminal domain of the Hfq protein with RNA molecules and a DNA fragment involved in the replication regulation of the ColE1 plasmid. The results of this research were described in the last experimental work included in my doctoral dissertation (**Appl. Sci. 2022; 12: 2639**). Due to the high molecular weight of the complexes formed between nucleic acids and the amyloid form of the protein, it is difficult to analyze solely by a gel shift assay the complexes formed, as they all migrate at the same position in the gel. In addition, precise kinetics measurements are not possible using a gel shift assay. Therefore, SRCD was used in a synchrotron-based biophysical approach to probe the interaction of the *E. coli* Hfq C-terminal amyloid region with nucleic acids involved in the control of ColE1-like plasmid replication. It was observed that this C-terminal region of Hfq has an unexpected and significant effect on the annealing of nucleic acids involved in this process and, more importantly, on their alignment. The functional consequences of this

newly discovered property of the Hfq amyloid region are important for understanding the biological importance of Hfq in the ColE1-type plasmid replication and antibiotic resistance of bacteria. In particular, the obtained results allowed to propose the hypothesis that Hfq plays a negative role in regulating the replication of the ColE1 plasmid by stimulating the interactions between RNA I and RNA II molecules, inhibiting the formation of an RNA II-DNA hybrid, which in turn reduces the efficiency of plasmid replication initiation. The decrease in the efficiency of transformation of *hfq* mutants with ColE1 plasmids may therefore result not from inhibition of replication of plasmid molecules in the absence of the Hfq protein, but rather from their excessive replication intensity (so called runaway replication of plasmids), leading to a decrease in the viability of bacterial cells that took up plasmid DNA molecules.

In conclusion, my research described in this dissertation led to the following key conclusions regarding the role of the Hfq protein in regulating the replication of ColE1-like plasmids:

1. The Hfq protein, and in particular its C-terminal domain, changes the mechanical properties and topology of the DNA double helix, which may be important in the regulation of DNA replication.

2. The efficiency of transformation of *hfq* mutants with ColE1-type plasmids is reduced, and interactions between Hfq and Rom in the regulation of replication of these plasmids are possible.

3. By stimulating the interaction between RNA I and RNA II, the Hfq protein, and in particular its C-terminal domain, can act as a negative regulator of the initiation of replication of ColE1-like plasmids.

4. Under the conditions of inactivity of the Hfq protein, the ColE1-like plasmids may over replicate in *E. coli* cells (so called runaway replication of plasmids), which leads to reduced viability of the bacteria.