CRISPR-Cas in *Escherichia coli*: regulation by H-NS, LeuO and temperature

**Abstract**

CRISPR-Cas adaptive immune systems are present in many bacteria and archaea and provide protection against invading DNA such as phages and plasmids. These systems are very versatile and complex in their gene composition and genomic architecture. CRISPR-Cas systems are classified into 2 classes, 6 types and 33 subtypes although this number is not definitive and the research is ongoing. All CRISPR-Cas systems have been thoroughly investigated in order to better understand the mechanism of CRISPR immunity enabling its use as a tool in genome editing and other biotechnological applications. However, regulation of the CRISPR-Cas system is also very complex and still not fully understood; it must provide optimal protection without introducing harmful consequences to the host. In this review, we give an overview on the regulation of the CRISPR-Cas system Class 1 Type I-E in *Escherichia coli* with the emphasis on the role of temperature in the regulation of the CRISPR-Cas activity and the interplay of the key regulators H-NS and StpA repressors and LeuO antirepressor in regulation of cas gene expression and HtpG chaperone in maintaining functional levels of Cas3.

**INTRODUCTION**

Clustered Regularly Interspaced Short Palindromic Repeats (CRISPR) and CRISPR-associated (Cas) proteins constitute the CRISPR-Cas immune systems, which show remarkable variability in gene composition, gene sequences, genomic architecture and activity (regulation). It is present in most archaea (~90%) and many bacteria (~40%). According to several criteria that take into account the signature *cas* genes, phylogeny of *cas1*, the organization of the genes in the CRISPR-Cas loci and in-depth sequence and structural analysis of the effector complexes (1-3), CRISPR-Cas systems are classified into two distinct classes (Class 1 and Class 2), which differ in the organization of the effector complex (1, 2). Class 1 CRISPR-Cas systems include types I, III, and IV, which utilize a multiprotein effector complex, while Class 2 CRISPR-Cas systems include types II, V and VI, which have a much simpler organization and utilize a single multifunctional effector protein (3). Of all identified CRISPR-Cas loci, Class 2 systems are less common than Class 1, possibly because they are less efficient in destroying phages, but are more easily transformed into numerous tools in molecular biology (4). Class 1 systems are widespread in both archaea and bacteria, while Class 2 occur almost exclusively in bacteria (5, 6). The Type I systems contain seven subtypes (I-A, I-B, I-C, I-D, I-E, I-F and I-U), which are defined by the presence of the multiprotein crRNA-effector complex (Cascade) and the Cas3 signature protein (1). Cas3 protein degrades invading DNA, using its helicase and nuclelease activities (reviewed in (7)), after...
its recognition by the effector complex. The Type I-E in *E. coli* is composed of a single operon containing *cas1* (*cas8e*), *cas2*, *cas7*, *cas5*, *cas6e*, *cas1* and *cas2* genes, the separate *cas3* gene, and the CRISPR 1 array (8).

Bacteria become resistant (immune) to invading DNA by inserting a small DNA fragment into the CRISPR array that consists of repeats interspaced with variable “spacer” sequences. Spacers originate from the invading DNA where they are called “protospacers”, and the process of acquiring new spacers is called adaptation (9, 10). Spacers allow bacteria to memorize and recognize the same invading DNA in the next infection. Through the process called expression and maturation, the CRISPR array is transcribed and processed to generate crRNAs (CRISPR-RNA) (11). Each crRNA is composed of a partial repeat(s) and a single spacer sequence. crRNA then assembles with Cas effector proteins into a ribonucleoprotein complex Cascade (CRISPR-associated complex for antiviral defense) that detects the complementary sequence in the invading DNA. An R-loop is formed when crRNA base pairs with the target strand of the invading DNA. This induces conformational change in the Cascade complex that recruits endonuclease Cas3 (12) and triggers degradation of target DNA in the process called interference (reviewed in (13)).

The activity of CRISPR-Cas systems is individual and depends on the bacterial species. Curiously, the expression of the CRISPR-Cas system in wild type *E. coli* is considered silent under laboratory growth conditions (14, 15). However, regulation of the CRISPR-Cas in *E. coli* is very complex - mediated at transcriptional and posttranslational level and strongly affected by the temperature of incubation and the cellular metabolism. The involvement of non-Cas protein HtpG and temperature as regulators makes CRISPR-Cas Type I-E in *E. coli* a unique and important model with which to study CRISPR-Cas regulation.

**CRISPR-Cas promoters and transcriptional regulation**

The components of the CRISPR-Cas system in *E. coli* are regulated by the three independent promoters Peas3, Peas8e and Perispr1. The role of the fourth, anti-Peas, is currently unknown, but it is assumed to have a role in transcription termination or prevention of the *cas3* mRNA translation (14). anti-Peas and Peas8e promoters are located in the intergenic region *ygiL-ygiB* (IGLB), or *cel1/casA-cas3*, but on opposite strands. Peas8e controls expression of the entire *cas* operon, Peas3 controls expression of the *cas3* gene, and Perispr1 controls pre-crRNA transcription (14). Peas3 and Perispr1 are partially repressed while Peas8e is fully repressed by the histone-like nucleoid structuring (H-NS) protein (14–16), an abundant nucleoid-associated protein (NAP). H-NS is a global regulator in many Gram-negative bacteria where it represses many genes involved in stress responses or changes in environmental conditions such as temperature, osmolarity, changes in pH or oxygen (17). In most cases, H-NS forms nucleoprotein complexes by spreading along AT-rich DNA and sterically hinders RNA-polymerase open complex formation (18), but it can also modulate expression of certain genes post-transcriptionally (17). Because H-NS binds onto AT-rich sequences such as those found in phages and plasmids (19), it has been proposed that horizontally acquired genes could titrate H-NS from promoters as an initial, passive step in anti-H-NS silencing. Indeed, the H-NS parologue Sfh (*Shigella flexneri* H-NS-like) protein encoded by plasmids R27 and pHCMI of *Salmonella Typhimurium* has been found to supplement the depleted population of temporarily decreased H-NS and enhance the fitness of bacteria when acquiring conjugative plasmids (20). Sfh protein also supports acquisition of plasmids by silencing CRISPR transcription (21).

The most important step in the activation of the CRISPR-Cas immunity system in *E. coli* is the expression of Cascade proteins. This is because the Cascade complex mediates the maturation of the crRNA by Cas6e endonuclease, target recognition by crRNA and target degradation by recruiting Cas3 (12, 22). A highly likely candidate for the relief of H-NS repression after titration of H-NS is the H-NS antagonist LeuO (transcriptional activator for *leuABCD* operon). The tetrameric DNA-binding protein LeuO is a transcription factor from the LysR family (23). LeuO is a pleiotropic positive regulator of many loci in Gram-negative bacteria, and LeuO and H-NS overlap in co-regulation of many target genes (78 % in *E. coli* and 40 % in *Salmonella enterica*) where LeuO acts as an antipressor (23, 24). LeuO positively regulates its own expression (25), and is also repressed by H-NS and StpA in both *E. coli* and *S. enterica* (26–28). StpA protein (suppressor of td phenotype A) is a nucleoid-associated multifunctional protein of similar size and amino acid identity as H-NS (29). It acts as a transcriptional repressor (30) and an RNA chaperone (31). Antagonistic regulators LeuO and H-NS have been recently posited to act as a genetic switch. The competition between H-NS and LeuO for access to the same DNA binding sites provides the basis of this switch (21). H-NS is at an advantage, since it is present at constant levels throughout the growth cycle (21), while the amounts of the LeuO protein will depend on its transient expression, which in turn depends on many conditions. Signals known to induce LeuO expression moderately are guanosine tetraphosphate (ppGpp) or amino acid starvation and stationary growth phase (24, 32, 33). However, neither increased transcription from Peas8e nor the formation of mature crRNA has been detected in cells with induced stringent response either by amino acid starvation or serine hydroxamate (15). Thus, it seems that the signals that induce LeuO or decrease H-NS are stochastic which makes the outcome of this competition different in each cell within the bacterial population (21).
Regulation of the Cascade complex levels

In addition to H-NS, transcription from the Pcas8e promoter is repressed by the cAMP-dependent DNA-binding protein CRP (34) and the H-NS paralogue StpA (35). CRP binds to a site that overlaps with a LeuO binding site, and makes CRISPR transcription dependent on glucose levels (34). StpA binds to the upstream region of the promoter Pcas8e and binds with a higher DNA binding affinity than H-NS in vitro (14). However, the silencing of CRISPR-Cas is mediated by H-NS, not by StpA in vitro (14). Since StpA is also repressed by H-NS (actually stpA and hns genes inhibit each other; (27), it is expressed only in cells lacking H-NS (36). StpA can substitute for H-NS in one third of H-NS repressed genes acting as a molecular backup (36–38). StpA is rapidly degraded in the absence of H-NS by the Lon protease (39).

We found recently that StpA partially represses transcription of the cas operon (genes cas5, cas1, cas2) in Δhns cells. Complete derepression was observed in Δhns ΔstpA cells or wt cells overexpressing LeuO; derepression was about 4-5-fold higher than in Δhns cells and about 250-fold higher than in wt cells. However, reduced transcription of Cascade genes in Δhns cells did not prevent maximal resistance to λvir phage (at low incubation temperature, i.e. 30 °C). The reason could be that sufficient amounts of Cascade complexes are synthesized even with modest transcript levels (35). Another reason may be that only few Cascade complexes are required to establish effective immunity thanks to the very fast scanning ability of DNA sequences (30 ms) and rapid movement between the nucleoid and the cytoplasm (40). In E. coli, about 20 molecules of Cascade complexes are estimated to provide about 50 % protection against the target establishment (interference levels), but the best chance to destroy the target DNA is before replication is initiated, i.e. when the target is present in a single copy. After that, the chance to destroy multiple copies of the target with the CRISPR-Cas system becomes very unlikely. Thus, a kinetic arms race takes place between the Cascade complex, which must find the target, and the invader before it gets permanently established (40).

Another layer of complexity of the system is the fact that destruction of the phage DNA in E. coli is strongly influenced by the temperature of incubation (16, 35). The Cas3 protein plays a critical role in this phenomenon (16).

Regulation of the Cas3 protein levels

In general, regulation of cas3 expression is much less understood than that of cas operon. The cas3 gene is partially negatively regulated by H-NS, and positively co-regulated by cAMP-dependent DNA-binding protein CRP and the glycine cleavage system (16, 41). Increased transcription of cas3 was observed in stationary Δhns cells (16), and in late log phase cells (OD600~1.0) incubated with glycine in minimal media or in LB with elevated levels of cAMP (41). Discovery of a positive role of CRP in regulation of cas3 expression is in contrast with its negative role in regulation of cas operon (34). This discrepancy is explained by different cell densities/growth phases of analyzed bacteria (41). Despite increased transcription of cas3 from the Pcas9 promoter in Δhns cells with introduced anti-lambda spacers, cells remained sensitive to λvir phage infection at 37 °C unless Cas3 was overexpressed from the plasmid or expressed from inducible chromosomal promoters. In addition, previous research found that the high-temperature protein G (HtpG) chaperone increases the stability and functional levels of Cas3 for efficient CRISPR-Cas interference (42). Taken together, these results suggested that the amounts of Cas3 protein are limiting at 37 °C (16).

Posttranslational regulation

HtpG belongs to the Hsp90 family of highly conserved ATP-dependent molecular chaperones that fold and remodel proteins. Hsp90 also functions as a “holdase” (prevents aggregation of client proteins) and is not essential in E. coli (43). Initial research on HtpG-Cas3 interaction found that the CRISPR system lost its suicidal activity against λ prophage and its ability to provide immunity from lysogenization at 32 °C in ΔhspG cells (42). Overexpression of HtpG or Cas3 restored CRISPR activity in lysogenization at 32 °C, but the interaction of HtpG-Cas3 in resistance to λvir phage infection at 37 °C is more complex. That is, overexpression of Cas3 could restore resistance to λvir phage infection at 30 °C without HtpG, in ΔhspGΔhns cells with anti-lambda spacers, but not at 37 °C where overexpression of Cas3 could restore resistance to λvir phage infection only in htpG Δhns cells (16). This corroborates the observation that the Cas3 protein requires HtpG at 37 °C, but the precise role of HtpG in this process is currently unknown. Curiously, the requirement for HtpG in maintaining Cas3 protein levels has been observed only in E. coli (37). In other bacteria, Hsp90 has been shown to be essential under specific conditions – at higher temperatures, for cold adaptation, swarming ability, biofilm formation, resistance to oxidative stress and participation in virulence (reviewed in (43)).

Regulation by temperature, supercoiling and growth phase

Many of the target genes that are negatively regulated by H-NS are regulated by environmental factors such as osmolarity, temperature, anaerobic stress, oxidative stress, pH or growth phase. These factors can induce changes in DNA topology (supercoiling) that can affect gene expression in response to those inducing factors (44). Therefore, it was proposed that the efficiency of CRISPR-Cas immunity can be linked to cellular metabolism via variable DNA supercoiling and metabolic flux (21). Negative supercoiling is known to stimulate R-loop formation, target recognition and spacer integration, which are more typical in the exponential phase of growth (10, 22, 45). There-
Therefore, bacteria in an exponential growth phase are more likely to resist infection by the CRISPR-Cas immune system than bacteria in the stationary growth phase, which have more relaxed DNA (21). Nevertheless, activation of cas3 by cAMP or H-NS inactivation was found in stationary bacteria as explained previously. This discrepancy can be explained by the linking number of bacterial DNA, which can change due to environmental stresses such as temperature (46). Temperature can also induce changes in the expression of certain genes or replication of the host or invading DNA (47). Indeed, the Δhns mutation was shown to strongly reduce plasmid and chromosomal copy number at 30 °C (48, 49). This effect was suggested to indirectly help CRISPR-Cas immunity to destroy invading DNA before replication is initiated (48).

In Pseudomonas aeruginosa, low temperature of incubation was shown to promote CRISPR-Cas immunity. Specifically, CRISPR adaptation was increased as a consequence of the accumulation of CRISPR-Cas complexes and slower growth, which gave more time for invading DNA to be destroyed (50). Temperature is also involved in the regulation of H-NS and StpA. H-NS is itself cold-shock regulated, being stimulated 3-fold by the cold-shock protein A (CspA) (51), while transcription of stpA from its native promoter is 2-fold decreased at 30 °C compared to 37 °C (36). LeuO is not temperature-regulated in wt cells (28), but due to reduced expression levels of StpA at lower temperature and its instability (39), it can be assumed that LeuO will have a stronger effect on activation of the CRISPR-Cas immunity in Δhns cells at 30 °C (35).

**Concluding remarks**

Antagonistic proteins H-NS and LeuO repress and derepress cas genes, aided by additional regulators StpA and cAMP-CRP to ensure sufficient amounts of Cascade complexes are expressed for efficient CRISPR-Cas-mediated immunity. On the other hand, regulation of Cas3 by H-NS and CRP does not seem to provide sufficient amounts of Cas3 protein to resist phage infection at 37 °C (16). At present, we do not know whether Cas3 protein is unstable or whether its activity is changed at 37 °C, but HtpG chaperone is required to maintain functional levels of Cas3 (42). Given that the incubation temperature determines whether cells will be resistant or sensitive to phage infection, we propose that incubation temperature is a switch that indirectly activates (turns ON) or inactivates (turns OFF) CRISPR-Cas immunity in E. coli (Figure 1). In fact, the temperature represents the Cas3 activity or inactivity; at 30 °C Cas3 is active while at 37 °C Cas3 is inactive (when expressed from its native promoter). It will be interesting to determine whether there are other, unknown, factor(s) involved in the regulation of the Cas3 activity or stability and why Cas3 activity is associated with the growth phase and temperature in E. coli.

**Acknowledgements:** This work was supported by the Croatian Science Foundation (Grant IP-2016-06-8861).

**Conflicts of Interest:** The authors declare no conflict of interest.
REFERENCES


29. ZHANG A, BELFORT M 1992 Nucleotide sequence of a newly-
identified Escherichia coli gene, stpA, encoding an H-NS-like pro-
30. WOLF T, JANZEN W, BLUM C, SCHNETZ K 2006 Differen-
tial dependence of StpA on H-NS in autoregulation of stpA and in condi-
32. FANG M, MAJUMDER A, TSAI K-J, WU H-Y 2000 ppGpp-
Dependent leuO expression in bacteria under stress. Biochem Bio-
36. SONDEN B, UHLIN BE 1996 Coordinated and differential ex-
37. SHI X, BENNETT GN 1994 Plasmids bearing hfq and the hns-
39. JOHANSSON J, UHLIN BE 1999 Differential protease-mediate-
40. VINK JNA, MARTENS KJA, VLOT M, MCKENZIE RE, AL-
MENDROS C, ESTRADA BONILLA B, BROCKEN DJW, HOHLEBEIN J, BROUNS SJ 2020 Direct visualization of native CRISPR target search in live bacteria reveals cascade DNA surveil-
41. YANG D, WANG Z, MA J, FU Q, WU L, WANG H, WANG S, YAN Y, SUN J 2020 Glycine cleavage system and cAMP recep-
tor protein co-regulate CRISPR/Cas3 expression to resist bacterio-
42. YOSEFI, GOREN MG, KIRO R, EDGAR R, QIMRON U 2011 High-temperature protein G is essential for activity of the Esche-
richia coli clustered regularly interspaced short palindromic repeats (CRISPR)/Cas system. Proc Natl Acad Sci USA 108(50): 20136–
20141. https://doi.org/10.1073/pnas.1113519108
45. SZCZELKUN MD, TIKHOMIROVA MS, SINKUNAS T, GA-
SIUNAS G, KARVELIS T, PSCHERA P, SIKSYNS V, SEIDEL R 2014 Direct observation of R-loop formation by single RNA-
46. GOLDSTEIN E, DRLICA K 1984 Regulation of bacterial DNA supercoiling: Plasmid linking numbers vary with growth tempera-
47. GUNASEKERA TS, CSONKA LN, PALIY O 2008 Genome-
wide transcriptional responses of Escherichia coli K-12 to continu-
48. RADOVIĆ M, CULO A, IVANČIĆ-BAĆE I 2019 Cas3-stim-
lated runaway replication of modified ColE1 plasmids in Esche-
49. ATLUNG T, HANSEN FG 2002 Effect of different concen-
51. TEANAALA, BRANDI A, FALCONI M, SPURIO R, PONCL, GUALERZI CO 1991 Identification of a cold shock transcrip-
tional enhancer of the genome-like gene encoding nucleoid pro-