High-temperature protein G is essential for activity of the *Escherichia coli* clustered regularly interspaced short palindromic repeats (CRISPR)/Cas system

Ido Yosef, Moran G. Goren, Ruth Kiro, Rotem Edgar, and Udi Qimron

*Department of Clinical Microbiology and Immunology, Sackler School of Medicine, Tel Aviv University, Tel Aviv 69978, Israel; and Divisions of Epidemiology, Molecular Epidemiology, and Antibiotic Resistance, Tel Aviv Sourasky Medical Center, Tel Aviv 64239, Israel

Edited by Sankar Adhya, National Institutes of Health, National Cancer Institute, Bethesda, MD, and approved October 28, 2011 (received for review August 18, 2011)

Prokaryotic DNA arrays arranged as clustered regularly interspaced short palindromic repeats (CRISPR), along with their associated proteins, provide prokaryotes with adaptive immunity by RNA-mediated targeting of alien DNA or RNA matching the sequences between the repeats. Here, we present a thorough screening system for the identification of bacterial proteins participating in immunity conferred by the *Escherichia coli* CRISPR system. We describe the identification of one such protein, high-temperature protein G (HtpG), a homolog of the eukaryotic chaperone heat-shock protein 90. We demonstrate that in the absence of htpG, the *E. coli* CRISPR system loses its suicidal activity against λ prophage and its ability to provide immunity from lysogenization. Transcomplementation of htpG restores CRISPR activity. We further show that inactive activity of the CRISPR system attributable to htpG deficiency can be suppressed by expression of Cas3, a protein that is essential for its activity. Accordingly, we also find that the steady-state level of overexpressed Cas3 is significantly enhanced following HtpG expression. We conclude that HtpG is a newly identified positive modulator of the CRISPR system that is essential for maintaining functional levels of Cas3.

**Results**

Positive Selection System for *E. coli* Genes Required for a Functional CRISPR System. To identify genes that are essential to the activity of the CRISPR system, we established a genetic screen that positively selects *E. coli* mutants with an inactive CRISPR system. The selection principle is based on the suicidal activity of the CRISPR system when *E. coli* λ lysogens are transformed with a plasmid encoding spacers against the λ phage (17). In such a system, transposon mutagenesis can be applied to identify viable bacterial clones with transposon disruptions of genes essential for CRISPR activity. We selected *E. coli* strain BW25113Δhns, in which the CRISPR system is constitutively active (9, 14, 17, 18), for the genetic screening. This strain was lysogenized with the λ phage, transscripts with the aid of an RNA molecule and the Cas protein, Cas9. H-NS and LeuO have been shown to regulate the CRISPR system in both *Escherichia coli* and *Salmonella typhi*, whereas LRP regulates the *S. typhi* CRISPR system (13–15). Other non-cas genes participating in the activity of the CRISPR system have not yet been identified in other bacteria. The identification of RNase III, as well as H-NS, LRP, and LeuO, suggests the existence of other non-cas gene products participating in the activity of the CRISPR system. Such gene products may have a direct role in CRISPR activity, such as DNA cleavage, crRNA processing, and mediating the encounter with incoming DNA, or they may have indirect supportive activities, such as proper folding of the involved proteins, energy supply for the system, regulation of the system, and bridging between protein machineries.

In this study, we established a system to search for genes whose products are required for activity of the *E. coli* CRISPR system. No such systematic search has ever been reported for any CRISPR system. The search is based on positive selection of *E. coli* colonies having disruptions in genes essential for activity of the CRISPR system. Our search coverage is thorough, consisting of 50,000 independent mutants covering over 10-fold the *E. coli* genome and identifying cas genes numerous times as essential for the CRISPR activity. We therefore believe that we have identified all the nonessential *E. coli* genes that are essential for the CRISPR activity. In particular, we identified high-temperature protein G (HtpG), a protein that is homologous to the eukaryotic chaperone heat-shock protein 90 (Hsp90) (16), as essential for maintaining functional levels of Cas3, a required protein for CRISPR interference with incoming DNA (3). A possible model for these interactions is discussed.

**Author contributions:** I.Y., M.G.G., R.E., and U.Q. designed research; I.Y., M.G.G., and R.E. performed research; R.E. contributed new reagents/analytic tools; I.Y., M.G.G., R.E., and U.Q. analyzed data; and U.Q. wrote the paper.

The authors declare no conflict of interest.

This article is a PNAS Direct Submission.

*1* Y. and M.G.G. contributed equally to this work.

*2* To whom correspondence should be addressed. E-mail: ehudq@post.tau.ac.il.

This article contains supporting information online at www.pnas.org/lookup/suppl/doi:10.1073/pnas.1113519108/-/DCSupplemental.
making the bacteria incompatible, in a CRISPR-dependent manner, with the plasmid pWUR478 encoding spacers against the λ phage (3, 17). In addition, the strain was genetically engineered to carry an arabinose-inducible T7-RNA polymerase (RNAP) to enhance expression of the CRISPR spacers, encoded downstream of a T7-promoter in the pWUR478 plasmid. When this strain was transformed with plasmid pWUR478 encoding anti-λ spacers, the transformation efficiency, compared with that of plasmid pWUR477 encoding control spacers, was reduced over 10,000-fold (Fig. 1), indicating that the spacers against the λ phage kill the cells in a CRISPR-dependent manner. To demonstrate further that the CRISPR system is responsible for the reduced transformation efficiency, we used an isogenic control strain, with an intact hns gene encoding the H-NS protein that represses CRISPR-system activity in vivo (9, 14). This strain showed similar transformation efficiency with pWUR478 and the control plasmid (Fig. 1). We concluded that the apparent decrease in transformation efficiency of strain BW25113Δhns by pWUR478 is attributable to cell death as a result of CRISPR-system activity. This significant difference in transformation efficiency enabled us to establish a robust system for the positive selection procedure by transforming pWUR478 into transposon-mutagenized bacteria and examining gene disruptions in the viable colonies.

Identification of E. coli Genes Essential for CRISPR-System Activity.

We next genetically engineered BW25113Δhns araB::T7-RNAP λc857-bla, a strain similar to the above but with a λ prophage having a different resistance marker, to harbor transposon insertions at various genetic locations, and we followed the selection scheme depicted in Fig. 2A. The generated transposon library represented ~50,000 independent events, providing over 10-fold coverage of the E. coli genome. This bacterial library was made competent, transformed with pWUR478 encoding anti-λ spacers, and selected on LB-agar plates supplemented with chloramphenicol, resistance to which is conferred by the plasmid. Bacteria able to harbor the plasmid were suspected of having an inactive CRISPR system because, as explained above, the lysogenic bacteria can only coexist with this plasmid in the absence of CRISPR activity. This procedure yielded 127 bacterial candidates that were further tested. We first eliminated transposition events in known CRISPR essential genes, the cas genes, from further analyses using multiplex PCR on these colonies. We used a mix of five oligonucleotides, three binding to the cas genes at 2- to 3.5-kb intervals, and two primers amplifying outward of the transposon (Fig. 2A). This procedure generates a PCR product when a transposon is inserted in one of the cas genes. Of the 127 candidates, 113 indeed harbored a transposon insertion in one of these genes, confirming the validity of the selection assay. A typical agarose gel showing PCR-amplified products from nine colonies selected for the ability to harbor pWUR478 is shown in Fig. 2B. Because cas1 and cas2 have been shown not to be required for CRISPR activity (3), and transposon insertion in cas1 cannot be excluded by the indicated PCR, we wanted to verify that, indeed, these genes do not contain any transposon insertions. As expected, PCR amplification of cas1-cas2 did not reveal any transposon insertion in these genes (Fig. 2B). These mutants were not examined further. The transposon insertion site was identified in the other 14 candidates by semirandom PCR amplification followed by sequencing, as described previously (19). Redundant candidates were eliminated from further analyses, and P1 transductions generating “clean deletions” were carried out for representatives of each disrupted gene. These reconstructed deletion mutants were then tested for transformation efficiency of pWUR478 compared with pWUR477, and only one showed significant restoration of transformation efficiency, suggesting that the CRISPR system is nonfunctional in the absence of this gene. Thus, the other originally isolated colonies probably carried secondary mutations (perhaps in the cas genes), or the transposon somehow modified the gene’s activity such that it could not be reconstituted by deletions.

Validating the Essentiality of the Identified Genes to CRISPR-System Functionality. The deleted gene showing consistent restoration of transformation efficiency represented four selected colonies having disruptions in htpG, two at position 495,128 and two at position 495,364 (Fig. 2). As described above, to verify that disruption of htpG is responsible for the lack of CRISPR activity, and that there are no secondary mutations causing this effect, we reconstructed a deletion of htpG in BW25113Δhns araB::T7-RNAP by P1 transduction of the htpG::kan cassette (20), and after removing the kan cassette, we lysogenized these cells with λc857-kan, yielding BW25113Δhns ΔhtpG araB::T7-RNAP λc857-kan. The newly constructed strain was tested for transformation efficiency of pWUR478 compared with pWUR477. As shown in Fig. 3A, the transformation efficiency of pWUR478, encoding anti-λ spacers, into the htpG deletion mutant was comparable to that of pWUR477, encoding control spacers, indicating that the CRISPR system does not function in the absence of htpG. To verify further the essentiality of htpG to CRISPR interference and to exclude, for example, polar effects of htpG deletion on other genes, we complemented the ΔhtpG mutant with a plasmid encoding htpG under an isopropyl-β-thiogalactopyranoside (IPTG)-inducible promoter. As shown in Fig. 3B, transformation efficiency of pAC-478, encoding anti-λ spacers, into the ΔhtpG mutant harboring the HtpG-encoding plasmid was over 1,000-fold lower than that with the control pAC-477, or in comparison to pAC-478 transformation into an ΔhtpG mutant harboring a control vector. These results indicated that expression of HtpG from a plasmid can restore
CRISPR activity to ΔhtpG cells, unequivocally confirming that htpG is essential for CRISPR-system activity.

We have previously shown that the CRISPR system is active against the establishment of lysogens (17). We wished to demonstrate that deficiency of htpG prevents this activity against lysogenization. We therefore transformed pWUR478 or pWUR477 into E. coli BW25113Δhns ΔhtpG araB::T7-RNAP or into BW25113Δhns araB::T7-RNAP as a control. These strains were tested for lysogenization efficiency by infecting them with λcI857-bla phage at 32 °C, conditions favoring the lysogenic cycle, and then determining the number of colonies acquiring antibiotic resistance encoded by the phage (Fig. 3C). In this assay as well, lack of htpG significantly decreased the CRISPR-dependent protection of the bacteria from lysogenization by over 1,000-fold compared with the isogenic strain carrying htpG (Fig. 3D). These results indicate that htpG is also essential for the E. coli CRISPR system’s prevention of phage lysogenization.

Expression of Cas3 Restores CRISPR Activity in the ΔhtpG Mutant. It is important to note that in both the lysogenization and transformation-efficiency assays described above for E. coli ΔhtpG, the size of the colonies harboring pWUR478 was smaller than that of those harboring pWUR477. These results led us to assume that CRISPR activity is not completely silent in the absence of htpG and that HtpG somehow modulates the CRISPR system’s functionality. HtpG is the bacterial homolog of the eukaryotic Hsp90 protein, which is a well-characterized chaperone (16). Because we established that HtpG is essential for CRISPR activity, we speculated that one or more of the Cas proteins are client proteins of this chaperone. Consequently, if HtpG is indeed essential for the accumulation of an active Cas protein, htpG deficiency should be suppressed by expression of this client protein in higher amounts than its endogenous expression. This speculation assumes that htpG deficiency results in a limiting amount of an active Cas protein and that this might be overcome by overexpressing this Cas protein from a plasmid. To test this possibility and to identify the
putative client protein, we transformed plasmids from the ASKA library (21) encoding each of the six proteins known to be essential for CRISPR interference in E. coli BW25113Δhns ΔhtpG araB::T7-RNAP. We then measured transformation efficiencies with a plasmid encoding anti-λ spacers, pAC-478, compared with a control plasmid, pAC-477. As shown in Fig. 4, expression of Cas3 was able to relieve the dependency of CRISPR activity on HtpG. In the absence of HtpG, increased expression of Cas3, but not of any other Cas protein, fully restored the CRISPR system’s ability to decrease the transformation efficiency of pAC-478 by over 1,000-fold. These results indicate that sufficient expression of Cas3 can overcome the absence of HtpG, suggesting that under htpG deficiency, Cas3 protein is the only Cas component whose level is limiting for CRISPR activity. The results suggest that HtpG either helps in recruiting Cas3 protein to the CRISPR interference stage or is involved in this protein’s folding and/or in preventing its aggregation, thereby maintaining functional levels of Cas3.

HtpG Expression Significantly Increases Levels of Overexpressed Cas3. Finally, if HtpG is indeed essential for maintaining functional levels of Cas3, its expression should support higher accumulation of Cas3. We first tested this possibility in vivo by transforming cas3 fused to GFP into E. coli BW25113ΔhtpG with either a plasmid encoding htpG or a control plasmid. As negative controls, we transformed the other cas genes required for CRISPR interference, fused to GFP, with either a plasmid encoding HtpG or a control plasmid. These bacteria were grown overnight on LB-agar plates supplemented with appropriate antibiotics, with or without l-arabinose for HtpG induction. To detect the level of fluorescence proportional to the levels of the GFP-fused proteins, plates were imaged and quantified as described in Materials and Methods. The measured fluorescence intensities indicated that the level of Cas3 increases over threefold when HtpG is expressed (Fig. 5 A and B). The increase in Cas3 level was only observed when HtpG expression was induced by l-arabinose, further validating that the increase in Cas3 levels is HtpG-dependent. The levels of the other tested Cas-GFP proteins were unaffected by the expression of HtpG, indicating that, of all the known CRISPR-effector proteins, HtpG specifically increases the steady-state level of Cas3-GFP in the bacterial cells. To confirm this in vivo observation further, we used Cas3 without GFP. Cas3 was overexpressed from a plasmid in BW25113ΔhtpG with or without HtpG, and its protein level was analyzed. As shown by SDS-gel electrophoresis of total cell extracts, coexpression of HtpG and Cas3 resulted, in this case as well, in greater than threefold increase in the level of Cas3 compared with its level without HtpG. The identity of Cas3 in the gel was confirmed by Western blot analysis against the His-tag preceding this protein (Fig. 5C). These findings confirmed that HtpG increases the steady-state level of Cas3. Taken together, our results show that HtpG’s essentiality to CRISPR-system activity is mediated by its effect on Cas3.

Discussion

We carried out a thorough experimental search for non-cas genes essential for the activity of the E. coli CRISPR system. This search yielded the htpG gene, which was confirmed to be essential for the suicidal activity of plasmid-encoded spacers against λ prophage, as well as for protection against lysogenization by λ phage. For the CRISPR system to interfere efficiently with invading DNA, it must go through most, if not all, of the following steps: (i) recognize alien DNA entry, (ii) induce transcription of cas genes, (iii) correctly fold or modify the translated Cas proteins, and (iv) process the crRNA and mediate binding of the processed crRNA-Cas proteins to the invading DNA/RNA, resulting in (v) cleavage
and destruction of the target DNA/RNA. Some of these steps, such as crRNA maturation by the CasABCDE proteins in the E. coli system, are fairly well understood (3), whereas others, such as sensing DNA entry, are poorly understood (22). Our screen to identify genes participating in the above steps contributes an additional piece of information to one of the interference steps in E. coli by establishing an important role for HtpG in maintaining levels of Cas3, a crucial component in CRISPR interference.

HtpG, a homolog of the ubiquitous eukaryotic Hsp90, is a protein that has been conserved from bacteria to humans (23), suggesting it might play a role in prokaryotic CRISPR systems other than the E. coli system as well. The precise role of the E. coli HtpG has remained elusive, compared with our detailed knowledge of its eukaryotic homolog. It is ubiquitous in E. coli, and its level is elevated following heat stress (hence, its name). HtpG-deletion mutants suffer from slow growth at high temperatures, as well as slight protein aggregation (24). HtpG has been shown to be involved with two client proteins, the ribosomal protein L2 and a linker polypeptide in the phycobilisome of Synechococcus elongatus, in an activity that characterizes it as a molecular chaperone (25–27). Recently, HtpG has also been shown to promote reactivation of heat-denatured luciferase in a reaction that requires the activity of another E. coli chaperone, DnaK (26).

We demonstrate that overexpression of Cas3 can relieve the dependency of CRISPR-system activity on HtpG, indicating that Cas3 is a limiting factor for this activity in the absence of HtpG. Furthermore, we show that coexpression of HtpG with Cas3 increases steady-state levels of the latter. Our results are not the first to suggest that HtpG is involved in stabilizing large protein complexes, and they are in line with a recent study showing that HtpG is required to stabilize the S. elongatus phycobilisome, a large protein complex, by specifically interacting with some of its components (27). Thus, in addition to the unique role of HtpG in the activity of the CRISPR system, we shed further light on its role as a chaperone by identifying its stabilization of another putative client protein.

Analyses of bacterial databases to identify a direct link between co-occurrence of Cas3 and HtpG are complicated by the lack of a straightforward way of identifying functional vs. nonfunctional CRISPR systems, because such co-occurrence is predicted based on the functionality of both proteins together. Nevertheless, basic analysis of bacterial databases using EcoCyc (28) indicates that of 33 E. coli strains, only 2 lack an htpG ortholog, and these 2 also lack cas3. All 7 examined Salmonella strains carry htpG, and 5 of them also encode cas3. Furthermore, all examined Mycoplasma and Synechococcus strains and most of the Streptococcus strains lack both htpG and cas3. On the other hand, all the examined Lactobacillus strains encode cas3 but lack htpG. In addition, ~50% of the archaeca encode cas3 (29), but almost all of them lack htpG (30). This basic analysis indicates that co-occurrence of Cas3 and HtpG are complicated by the lack of other putative client protein.

To summarize, we identified HtpG as a non-Cas protein involved in CRISPR activity using a unique screening method tailored to the CRISPR system. In future studies, we will refine and modify our screening method to identify additional regulatory proteins in this system. We believe that regulatory genes thus identified and characterized will shed more light on the regulation of this intriguing system.

Materials and Methods

Reagents, Strains, Plasmids, and Construction Procedures. Reagents, strains, plasmids, and construction of strains and plasmids are detailed in SI Materials and Methods.

CRISPR-Dependent Restriction of Transformation. E. coli BW25113 Δhns araB::T7-RNAp λ::857-kan was diluted 1:50 from an overnight culture and aerated at 32 °C in LB medium containing 25 μg/mL kanamycin to an OD600 of
We thank John van der Oost and Stan Brouns for their contributions to this work. A widespread system that provides such resistance is encoded by a CRISPR-Cas system that is functional in a wide variety of bacteria. We also thank Nir Osherov for critical reading of the manuscript and Camille Vainstein for professional language editing. This research was supported by the Israel Science Foundation (Grant 611/10), the Binational Science Foundation (Grant 2009/218), the German-Israeli Foundation (Grant 2061/2009), and Marie Curie International Reintegration Grants (PIRG-GA-2009-256340 and GA-2010-266717).

**ACKNOWLEDGMENTS.** We thank John van der Oost and Stan Brouns for providing the plasmids encoding the CRISPR spacers, William Metcalf for providing the plasmid for generation of transposon insertions, David Ze’evi for technical help, and Lynn C. Thomas for providing the scbS7-Kan phage. We also thank Nir Osherov for critical reading of the manuscript and Camille Vainstein for professional language editing. This research was supported by the Israel Science Foundation (Grant 611/10), the Binational Science Foundation (Grant 2009/218), the German-Israeli Foundation (Grant 2061/2009), and Marie Curie International Reintegration Grants (PIRG-GA-2009-256340 and GA-2010-266717).


Quantification of Cas Proteins Following Coexpression with HtpG. To examine the steady-state level of Cas3 following HtpG overexpression, BW25113ΔhtpG and either pBAD18 (control) or pBAD-HtpG, along with pCA24N-Cas3-GFP [ASKA library (21)] or other specified control plasmids from the ASKA library, were cotransformed into E. coli. Cells were grown overnight in LB supplemented with the appropriate antibiotics at 37 °C. Aliquots of 0.1 mL were transferred into a 96-well microtiter plate, and the bacteria were then replicated on an agar plate with the appropriate antibiotics, with or without 0.2% l-arginine for induction of HtpG expression. The pCA24N-based plasmids contain a leaky lac promoter, and there is a sufficient amount of GFP for monitoring purposes, even without IPTG induction (21). After overnight growth on plates, the replicated bacteria were scanned for GFP fluorescence using a Typhoon 9400 (GE Healthcare) (488-nm excitation wavelength, 520-nm emission wavelength). The presented image was digitally optimized for background omission and uniform expression levels. Results were quantified using ImageJ-based (National Institutes of Health) densitometer. To analyze protein quantities by gel electrophoresis, E. coli BW25113ΔhtpG was transformed with either pBAD18 or pBAD-HtpG along with pCA24N-Cas3 (no GFP, from ASKA library). Cells were grown to the logarithmic phase and induced with 0.2% l-arginine. After 20 min of induction, the cultures were divided and IPTG was added to half of the sample. After an additional 1.5 h of growth, the cells were transferred to ice. Cell pellets (5 × 10^6 cells) were suspended in 1 mL of buffer A (1 mM EDTA, 50 mM Tris-HCl (pH 8), 100 mM NaCl) containing protease inhibitor mixture (Roche). Protein extracts were prepared by three cycles of brief 10-s sonication at 20-s intervals on ice. Equal amounts of total protein (20 μg) were separated on a 10% w/v polyacrylamide gel and stained with Coomassie blue. Varied conditions, such as using anti-His-tag antibody (Sigma) according to the manufacturer’s instructions.