An Active Immune Defense with a Minimal CRISPR (Clustered Regularly Interspaced Short Palindromic Repeats) RNA and without the Cas6 Protein*

Received for publication, October 10, 2014, and in revised form, December 12, 2014. Published, JBC Papers in Press, December 15, 2014, DOI 10.1074/jbc.M114.617506

Lisa-Katharina Maier1,2, Aris-Edda Stachler1, Sita J. Saunders3, Rolf Backofen4,5, and Anita Marchfelder1,2

From the 1Department of Biology II, Ulm University, 89069 Ulm, the 2Bioinformatics Group, Department of Computer Science, Albert-Ludwigs-University Freiburg, Georges-Köhler-Allee 106, 79110 Freiburg, and the 3BIOSS Centre for Biological Signalling Studies, Cluster of Excellence, Albert-Ludwigs-University Freiburg, Schanzlestrasse 18, 79104 Freiburg, Germany

The prokaryotic immune system CRISPR-Cas (clustered regularly interspaced short palindromic repeats–CRISPR-associated) is a defense system that protects prokaryotes against foreign DNA. The short CRISPR RNAs (crRNAs) are central components of this immune system. In CRISPR-Cas systems type I and III, crRNAs are generated by the endonuclease Cas6. We developed a Cas6b-independent crRNA maturation pathway for the Haloflex type I-B system in vivo that expresses a functional crRNA, which we termed independently generated crRNA (icrRNA). The icrRNA is effective in triggering degradation of an invader plasmid carrying the matching protospacer sequence. The Cas6b-independent maturation of the icrRNA allowed mutation of the repeat sequence without interfering with signals important for Cas6b processing. We generated 23 variants of the icrRNA and analyzed them for activity in the interference reaction. icRNAs with deletions or mutations of the 3’ handle are still active in triggering an interference reaction. The complete 3’ handle can be removed without loss of activity. However, manipulations of the 5’ handle mostly led to loss of interference activity. Furthermore, we could show that in the presence of an icrRNA a strain without Cas6b (Δcas6b) is still active in interference.

Background: CRISPR RNAs (crRNAs) are generated by Cas6b in type I-B systems. They are essential for the interference reaction.

Results: An icrRNA is generated independently from Cas6b and functions like a crRNA.

Conclusion: In the presence of an icrRNA, Cas6b is not required for the interference reaction.

Significance: This setup allows the Cas6b-independent generation of icrRNAs and thereby interference without Cas6b.

The prokaryotic immune system CRISPR-Cas (clustered regularly interspaced short palindromic repeats–CRISPR-associated) is the defense mechanism that protects prokaryotes against foreign DNA. The short CRISPR RNAs (crRNAs) are central components of this immune system. In CRISPR-Cas systems, type I and III, crRNAs are generated by the endonuclease Cas6. We developed a Cas6b-independent crRNA maturation pathway for the Haloflex type I-B system in vivo that expresses a functional crRNA, which we termed independently generated crRNA (icrRNA). The icrRNA is effective in triggering degradation of an invader plasmid carrying the matching protospacer sequence. The Cas6b-independent maturation of the icrRNA allowed mutation of the repeat sequence without interfering with signals important for Cas6b processing. We generated 23 variants of the icrRNA and analyzed them for activity in the interference reaction. icRNAs with deletions or mutations of the 3’ handle are still active in triggering an interference reaction. The complete 3’ handle can be removed without loss of activity. However, manipulations of the 5’ handle mostly led to loss of interference activity. Furthermore, we could show that in the presence of an icrRNA a strain without Cas6b (Δcas6b) is still active in interference.

*This work was supported by the German Research Council (Deutsche Forschungsgemeinschaft), through the research group FOR 1680 “Unravelling the prokaryotic immune system” (grant DFG MA1538/17-1 and BA2168/5-1).

† Both authors contributed equally to this work.

‡ To whom correspondence should be addressed: Biology II, Ulm University, 89069 Ulm, Germany. Tel.: 49-731-5022658; Fax: 49-731-5012-22658; E-mail: anita.marchfelder@uni-ulm.de.

§ The abbreviations used are: CRISPR, clustered regularly interspaced short palindromic repeat; Cas, CRISPR-associated; crRNA, CRISPR RNA; icrRNA, independently generated crRNA; nt, nucleotide; PAM, protospacer adjacent motif.

This is an Open Access article under the CC BY license.
is catalyzed by the Cas6 protein in CRISPR-Cas type I and type III systems. In some type I systems, Cas6 is part of the CRISPR-associated complex of antiviral defense (Cascade) (12) that consists of different Cas proteins depending on the subtype (2). In contrast, in the type III system Cas6 is a standalone endonuclease (13, 14). Processing by Cas6 within the repeat sequence directly yields the mature functional crRNA in types I-A, I-E, and I-F (9). The resulting crRNA consists of an eight-nucleotide repeat-derived 5′ handle, the invader-targeting spacer sequence, and the 3′ handle, which contains the remainder of the repeat sequence (Fig. 1A) (2). In some type I systems (I-E and I-F), the Cas6 proteins stay bound to the crRNA after processing. In type III systems, a second maturation step is observed after Cas6 processing, which shortens the crRNA 3′ end and sometimes removes the complete repeat sequence downstream of the spacer (14, 15).

The initial invader DNA recognition is governed by Watson-Crick base pairing with a 7–10-nt segment of the crRNA referred to as the “seed” sequence (16–19). The seed sequence is involved in initial pairing between crRNA and invader, and it allows rapid probing of different regions of cellular nucleic acids. If a perfect match between seed sequence and target DNA is found, the remainder of the spacer sequence of the crRNA base pairs with the invader DNA. In the type I-E system, the seed sequence is a seven-nucleotide-long noncontiguous sequence between the 5′ end of the crRNA spacer sequence and the invader (17). In the type I-B system, this seed sequence is slightly longer with 10 nucleotides (20). An additional prerequisite for the interference is the presence of the PAM sequence in the invader DNA (2).

Here, we investigate the function of Cas6 in the interference reaction and the essential requirements for the crRNA in the type I-B system of the archaeon Haloferax volcanii. H. volcanii contains only one CRISPR-Cas system (I-B) that consists of eight Cas proteins (Cas1 to Cas5, Cas6b, Cas7, and Cas8b) and three CRISPR RNA arrays (20). We could previously identify the PAM sequences for this system showing that six different PAMs are active in triggering degradation (21). The Haloferax I-B system has a Cascade-like complex, with Cas6b copurifying with the Cas5 and Cas7 proteins and the crRNA (22). It has been shown that the Cas6b protein is involved in crRNA maturation and that the crRNA 5′ handles are eight nucleotides long; however, different 3′ lengths have been reported (22).

We developed here a Cas6b-independent crRNA maturation pathway for the Haloferax type I-B system in vivo that expresses a functional crRNA, which we termed independently generated crRNA (icrRNA). The icrRNA is transcribed with flanking tRNA-like structures (so-called t-elements) that are processed by the tRNA processing enzymes RNase P and tRNase Z (23). The icrRNA is effective in triggering degradation of an invader plasmid carrying the matching protospacer sequence.

We show here that a minimal crRNA in the I-B system needs a seven-nucleotide 5′ handle and does not require a 3′ handle at all. In addition, we show that the Cas6b protein is not required for the interference reaction when an icrRNA is present. With the Cas6b-independent maturation pathway developed here, the first in vivo analysis of crRNA characteristics essential for the interference reaction was possible.

**EXPERIMENTAL PROCEDURES**

**Strains—** H. volcanii strains H119 (strains used are listed in Table 1), Δcas6 (ΔpyrE2, ΔleuB, ΔtrpA, and Δcas6) (22), and ΔC (ΔpyrE2, ΔleuB, ΔtrpA, and HVO_2,385,045–2,386,660::trpA)

---

**TABLE 1**

<table>
<thead>
<tr>
<th>Strains</th>
<th>Plasmids</th>
<th>Primers</th>
</tr>
</thead>
<tbody>
<tr>
<td>DH5α</td>
<td>F- 80bacZAM15 Δ(lacZYA-argF16)U169 recA1 endA1 hsdR17 (rk-, mk+) gVal- supE44 λ- thi-1 gyrA96 relA1</td>
<td>ite1el</td>
</tr>
<tr>
<td>H119</td>
<td>ΔpyrE2 ΔtrpA ΔleuB</td>
<td>del1</td>
</tr>
<tr>
<td>Δcas6</td>
<td>ΔpyrE2, ΔleuB, ΔtrpA, Δcas6</td>
<td>del2</td>
</tr>
<tr>
<td>ΔC</td>
<td>ΔpyrE2, ΔleuB, ΔtrpA, ΔHVO_2,385,045–2,386,660::trpA</td>
<td>del3</td>
</tr>
<tr>
<td></td>
<td></td>
<td>del4</td>
</tr>
<tr>
<td></td>
<td></td>
<td>C-Sp1</td>
</tr>
<tr>
<td></td>
<td></td>
<td>D-OmitteC</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Cdelup</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Cdelupi</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Cdeluo</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Cdeluo</td>
</tr>
<tr>
<td></td>
<td></td>
<td>ZYA-arg</td>
</tr>
</tbody>
</table>

---

**Note:**

The Cas proteins Cas6 and Cas8 are termed Cas6b and Cas8b in CRISPR-Cas Haloferax.
Prokaryotic Immune Defense with an icrRNA and without Cas6

VOLUME 290 • NUMBER 7 • FEBRUARY 13, 2015

4194 JOURNAL OF BIOLOGICAL CHEMISTRY

A. Sabag-Daigle and C. J. Daniels, manuscript in preparation.

The plasmids for expressing icrRNA (pTA409-telectRNA, pTA232-telectRNA, and telecrRNA variants in both vectors) were generated as follows (plasmids are listed in Table 1). The DNA fragment containing the crRNA or crRNA mutants flanked by t-elements were ordered from GeneArt® as plasmids pMA-RQ-telecRNA and pMA-telecRNA. Plasmids contained a synthetic Haloferax promoter,5 the crRNA, flanked by t-elements and a synthetic Haloferax terminator.5 Plasmids were digested with KpnI and BamHI to isolate the DNA fragment containing the complete insert. The resulting fragment was cloned into pTA409 (26) and pTA232 (27) (both digested with KpnI and BamHI). Four crRNA mutants were generated by inverse PCR on pMA-telecRNA using primer pairs (primer sequences are listed in Table 1) itelle1/del1, itelle1/del1, itelle1/del1, and itelle1/del1 to generate variant 13 (deletion of the last five nucleotides of the 3’ handle), 14 (deletion of the last 10 nucleotides of the 3’ handle), 15 (deletion of the last 15 nucleotides of the 3’ handle), and 16 (deletion of the last 20 nucleotides of the 3’ handle), respectively. In preparation for transformation, all plasmids were passaged through E. coli GM121 cells to avoid methylation. Haloferax cells were subsequently transformed using the polyethylene glycol method (27, 28).

Generation of a CRISPR Locus C Gene Deletion Strain (ΔC)—The deletion of the CRISPR locus C was achieved by using the pop-in/pop-out method as described previously (24, 25, 29). The region upstream of the gene for CRISPR locus C was PCR-amplified with flanking regions from the chromosomal DNA of H. volcanii strain H119 using primers Cdel1 and Ddel1 (containing the restriction site XbaI) and Cdeldoi (containing the restriction site EcoRV). The resulting 500-bp fragment was digested with EcoRV to insert the marker gene trpA (coding for tryptophan synthase A). The tryptophan marker trpA was amplified using plasmid pTA131 (digested with EcoRV and XbaI), yielding pTA131-Cup. Next, the region downstream of the locus C gene was amplified using primers Cdeldoi (containing the restriction site XbaI) and Cdelupi (containing the restriction site KpnI) and Cdeldoi (containing the restriction site KpnI) and Cdelupi (containing the restriction site KpnI) and Cdelupi (containing the restriction site KpnI) and Cdeldoi (containing the restriction site KpnI). The resulting 500-bp fragment was cloned into the plasmid pTA131-Cup (digested with EcoRV and XbaI), yielding plasmid pTA131-Cupdo. This plasmid was digested with EcoRV to insert the marker gene trpA (coding for tryptophan synthase A). The tryptophan marker trpA was amplified using plasmid pTA132 (27) as template and oligonucleotides TRP1/TRP2, and cloning of the trpA marker gene into the plasmid pTA131-Cupdo resulted in pTA131-CupdoTrp. Plasmids were passaged through E. coli GM121 to prevent methylation, and H. volcanii strain H119 was subsequently transformed with this construct to allow integration (pop-in) of the plasmid into the genome. The subsequent selection for loss of the pyrE2 marker by plating on 5-fluoroorotic acid revealed pop-out mutants. To confirm the removal of the gene for CRISPR locus C, we performed a Southern blot analysis. Chromosomal DNA was isolated from the wild type and potential locus C gene deletion mutants. Southern blot hybridization was performed as described (27), with the following modifications. 10 μg of SacII-digested DNA was separated on a 0.8% agarose gel and transferred to a nylon membrane (Hybond™-N, GE Healthcare). A 250-bp fragment of the downstream region of locus C was amplified using primers Cdeldoi and DOMitieC, and the fragment was radioactively labeled using [α-32P]dCTP and random primer kit Readprime™II (GE Healthcare) and subsequently used as a hybridization probe.

Plasmid Invader Tests—The invader plasmid constructs pTA352-PAM3Csp1 (30) and pTA409-PAM3Csp1 (16) were generated based on the Haloferax shuttle vectors pTA352 (pHV1, leuB) (31) and pTA409 (pHV1, pyrE2) (26), including spacer 1 of the CRISPR locus C (C1) and the PAM sequence TTC (PAM3) (16, 21). As a control reaction, Haloferax cells expressing the icrRNA (WT or mutants) were transformed with the vector without insert (pTA352 or pTA409). Plasmids were passaged through E. coli GM121 cells (to avoid methylation) and were then introduced into Haloferax cells using the PEG method (27, 28). To confirm the identification of a functional invader sequence, H. volcanii cells were transformed at least three times with the plasmid invader construct or the control vector. For plasmid invader tests, transformations with at least a 100-fold reduction in transformation rates are considered successful interference reactions (21, 32). High reductions in transformation rates provide evidence for high targeting efficiency of the crRNA analyzed.

Northern Blot Hybridization—Total RNA was isolated, unless stated otherwise, from exponentially growing H. volcanii cells as described (16). After separation of 10 μg of RNA (total RNA) on 8% denaturing gels, RNA molecules were transferred to nylon membranes (Hybond-N+, GE Healthcare) and incubated with oligonucleotides against the spacer 1 from locus C (primer C1). The primer was radioactively labeled at the 5’ end with [γ-32P]ATP and subsequently used for hybridization.

Investigation of icrRNAs—To determine the exact length and sequence of the crRNA, RNA was isolated from wild type Haloferax cells (H119) and strain ΔC × pTA232-telecRNA grown to an absorbance of 0.74. RNA was separated on 8% PAGE, and RNA ranging in size from 45 to 55 nucleotides (fraction 1) and from 60 to 75 nucleotides (fraction 2) was eluted and sent to vertis Biotechnologie AG for cDNA preparation and RNAseq analysis. The RNA samples were first treated with polynucleotide kinase and then poly(A)-tailed using poly(A) polymerase. Afterward, an RNA adapter was ligated to the 5’-monophosphate of the RNA. First-strand cDNA synthesis was performed using an oligo(dT)-adapter primer and the Moloney murine leukemia virus reverse transcriptase. The resulting cDNAs were PCR-amplified to about 10–30 ng/μl using a high fidelity DNA polymerase. The cDNAs were purified using the Agencourt AMPure XP kit (Beckman Coulter Genomics) and were analyzed by capillary electrophoresis. For Illumina sequencing, the cDNA samples were mixed in approximately equal amounts. An aliquot of the cDNA pool was analyzed by capillary electrophoresis. The primers used for PCR

5 A. Sabag-Daigle and C. J. Daniels, manuscript in preparation.
amplification were designed for TruSeq sequencing according to the instructions of Illumina.

RNAseq Mapping—First, original reads were trimmed according to their sequencing quality using the fastq_quality_trimmer program from the FASTX-Toolkit version 0.13 with the options “–t 13-Q 33.” The parameter Q is required due to the ASCII offset of 33 used for the quality scores in the Sanger format. The estimated probability that a base call is incorrect ($p > 0.05$) corresponds to quality values below 13 (33). Second, trimmed reads were mapped with Segemehl (34) version 0.1.3 with the options “–polyA –prime3/AGATC/–H11032 trimmed reads were mapped with Segemehl (34) version 0.1.3 with the options “–polyA –prime3/AGATC/GGAAGAGGCC-TGTAGGAAGGTGATCTCGGTGGTGCG-CGGTATCAT”.

This setting removes the poly(A) tail and the 3' Illumina sequencing adapter. The following percentages of the original reads were successfully mapped from each sample: 86% for S1 (wild type RNA fraction of 60–75 nt length), 74% for S2 (wild type RNA fraction of 45–55 nt length), 91% for S3 (icrRNA fraction of 60–75 nt length), and 81% for S4 (icrRNA fraction of 45–55 nt length). All samples had 20–40 million reads. Subsequent to mapping, alignments were filtered such that they had a maximum edit distance of 2, were located on the reverse strand (because CRISPR locus C is transcribed from the reverse strand), and matched uniquely to the genome. The filtering produced a clearer signal, but it did not change original profiles. To explore and display RNAseq results, we used the Integrative Genomics Viewer version 2.0.3 (35).

RESULTS

To determine the essential nucleotides of the crRNA for the interference and to investigate whether the Cas6b protein is required for the interference reaction, we established a Cas6b-independent crRNA generation in *H. volcanii*. Using this setup, we could study the effect of crRNA mutants on the interference reaction independently of the crRNA processing stage; thus, we captured crRNA characteristics that were specific to the interference reaction.

**Cas Protein-independent Generation of crRNAs**—We generated a plasmid that encodes the crRNA as well as t-elements directly up- and downstream of the crRNA (Fig. 1B). The crRNA is derived from the *Haloferax* CRISPR locus C and contains spacer 1 of this locus. The t-element is a tRNA-like structure that has been previously detected directly upstream of the 5S rRNA in *H. volcanii*, and it is processed by tRNase Z to generate the 5S rRNA 5' end (36, 37). Generally, t-elements are substrates for both tRNA-processing enzymes, the 5'-processing enzyme tRNase P, and the 3'-processing enzyme tRNase Z (36, 38). Processing of the t-elements up- and downstream of the crRNA should yield the mature icrRNA. We cloned the crRNA/t-element insert into the *Haloferax* vector pTA409 (26), yielding pTA409-telecrRNA. A *Haloferax* strain that has the CRISPR locus C deleted (strain ΔC) was generated to get a strain without the endogenous spacer 1 from locus C (see under “Experimental Procedures”). This strain was transformed with plasmid pTA409-telecrRNA yielding ΔC × pTA409-telecrRNA. Northern blot analysis showed that an icrRNA is generated with the same size as the crRNA made in the wild type strain (which generates the crRNA from the CRISPR locus C) (Fig. 1C). Thus, the icrRNA is efficiently generated from the plasmid. In addition, some shorter RNAs are visible, and these shorter crRNAs have also been reported earlier in wild type cells (22). Because the amount of icrRNA was rather low compared with the endogenous crRNA, we cloned the crRNA/t-element insert into a *Haloferax* vector with a higher copy number, pTA232 (27), yielding pTA232-telecrRNA. Northern analysis showed that a *Haloferax* ΔC strain transformed with pTA232-telecrRNA indeed generated higher amounts of icrRNA (Fig. 1C).

To confirm that processing of the icrRNA yielded exactly the same 5' and 3' ends as in the "natural" crRNA production, we isolated the two RNA fractions that contained the long crRNA of about 65 nucleotides (RNA fraction of 60–75 nucleotides in length isolated) and the shorter crRNA of about 51 nucleotides (RNA fraction of 45–55 nucleotides in length isolated) from wild type *Haloferax* cells and ΔC × pTA232-telecrRNA cells and analyzed them with RNAseq. The icrRNAs from the 60–75-nucleotide fraction (isolated from ΔC × pTA232-telecrRNA strain) have exactly the same 5' and 3' ends as the wild type crRNA (Fig. 2A). Thus, we could show that we can generate a mature icrRNA identical to the natural crRNA in *Haloferax* cells. In addition we could show that a slightly shorter icrRNA version with 49 nucleotides in length (icrRNA49) is also present (Fig. 2B). This shorter icrRNA49 has the same 5' end but a 17-nucleotide shorter 3' handle than icrRNA46.

The only difference between the natural crRNA and the icRNAs is the nature of the processing product end groups; the icRNA contains a 5'-phosphate group at the crRNA 5' end and a 3' hydroxyl group at the crRNA 3' end due to processing by tRNase P and tRNase Z (23, 39). This is in contrast to the observed end groups generated naturally by type I Cas6 processing as follows: a 5' hydroxyl group and 2'–3' cyclic phosphate (I-C and I-E) (13, 40, 41) or a noncyclic 3' phosphate (I-F) (18). However, we show here that the nature of the end group is not important for the interference reaction (see below). Taken together, we could successfully establish a Cas6b-independent crRNA maturation pathway.

**icRNAs Are Active in Interference**—To investigate whether the icRNA is active in interference, we challenged *Haloferax* strain ΔC expressing the icrRNA (from the high copy plasmid pTA232-telecrRNA) with an invader plasmid (21). The invader plasmid contains the protospacer sequence that matches the crRNA 5' end and a 3' hydroxyl group at the crRNA 3' end due to processing by tRNase P and tRNase Z (23, 39). This is in contrast to the observed end groups generated naturally by type I Cas6 processing as follows: a 5' hydroxyl group and 2'–3' cyclic phosphate (I-C and I-E) (13, 40, 41) or a noncyclic 3' phosphate (I-F) (18). However, we show here that the nature of the end group is not important for the interference reaction (see below). Taken together, we could successfully establish a Cas6b-independent crRNA maturation pathway.

**icRNAs Are Active in Interference**—To investigate whether the icRNA is active in interference, we challenged *Haloferax* strain ΔC expressing the icrRNA (from the high copy plasmid pTA232-telecrRNA) with an invader plasmid (21). The invader plasmid contains the protospacer sequence that matches the crRNA 5' end and a 3' hydroxyl group at the crRNA 3' end due to processing by tRNase P and tRNase Z (23, 39). This is in contrast to the observed end groups generated naturally by type I Cas6 processing as follows: a 5' hydroxyl group and 2'–3' cyclic phosphate (I-C and I-E) (13, 40, 41) or a noncyclic 3' phosphate (I-F) (18). However, we show here that the nature of the end group is not important for the interference reaction (see below). Taken together, we could successfully establish a Cas6b-independent crRNA maturation pathway.
reaction. Taken together, the icrRNAs can trigger the interference reaction and thus are fully functional crRNAs.

Cas6b Is Not Required for Interference in the Presence of icrRNAs—In the wild type situation Cas6b is required for crRNA production, and it is conceivable that it could also be required for the interference reaction, because it was shown to be part of Cascade in Haloferax (I-B system), E. coli (I-E), Pseudomonas aeruginosa (I-F), and Sulfolobus solfataricus (I-A) (18, 22, 42–46). By the Cas6b-independent generation of icrRNAs, we separated the role of Cas6b in crRNA processing from its function in the interference reaction. Using icrRNAs, we can now determine whether Cas6b is also important for the inter-

FIGURE 1. Natural crRNA of Haloferax and the icrRNA. A, crRNAs of Haloferax. Haloferax encodes three different CRISPR loci, P1, P2, and C, that have the same 30-nucleotide-long repeat sequences except for the first nucleotide of the 5’ handle (position −8 according to the nomenclature (42–44)) that is an A in P1, a U in P2, and a G in C. Thus, there are three types of crRNAs in Haloferax beginning with three different nucleotides. The mature crRNA contains an 8-nucleotide 5’ handle and a 22-nucleotide 3’ handle. Spacers are between 34 and 39 nucleotides long. Nucleotides in the 5’ handle are termed −8 to −1 (from the 5’ end of the 5’ handle) and nucleotides from the 3’ handle are termed +1 to +22 (42–44). B, maturation of the icrRNA. The pre-icrRNA contains the crRNA flanked by two t-elements. The crRNA is derived from CRISPR locus C containing spacer 1 from this locus. The t-elements are recognized and processed by RNase P and tRNase Z, generating the mature icrRNA of 66 nucleotides (icrRNA66). This icrRNA can be processed further to a 49-nucleotide-long icrRNA49 by still unknown RNases. C, maturation of the icrRNA in Haloferax cells. RNA was isolated from wild type cells (lane wt), Haloferax cells without the CRISPR locus C (lane ΔC), and ΔC cells with pTA409-telecrRNA (lane ΔC + in the left panel) and from ΔC cells with the high copy plasmid pTA232-telecrRNA (lane ΔC + in the right panel), respectively. After separation on 8% PAGE, the RNA was transferred to a membrane that was subsequently hybridized with a probe against the crRNA. The mature crRNA can be detected in wild type Haloferax cells but not in ΔC. Left panel, “low copy,” generation of icrRNAs from low copy plasmids. The mature icrRNA can be detected in ΔC transformed with the low copy plasmid pTA409-telecrRNA. Lane m, DNA size marker, sizes are given at the left in nucleotides. The icrRNAs are shown schematically at the right. Right panel, “high copy,” generation of icrRNAs from high copy plasmids. In lane ΔC +, the precursor of the icrRNA as well as the processing intermediates are visible. The long exposure (bottom right, “long”) shows that the shorter icrRNA of about 49 nucleotides is also present. Sizes of a DNA marker are given at the left in nucleotides. The precursor of the icrRNA, the intermediates, and the mature icrRNAs are shown schematically at the right.
Prokaryotic Immune Defense with an icrRNA and without Cas6

A comparison of Cas6b catalyzed crRNA generation (wt) and Cas6b independent crRNA production (icrRNA). RNAseq data from RNA fractions (sizes 60–75 nucleotides) isolated from wild type Haloferax cells (upper row “wt”) and ΔC × pTA232-telecrRNA (lower row “icrRNA”) were mapped to the CRISPR C locus. The icrRNA only comprises spacer 1, between repeats 1 and 2. The numbers to the right of each row reflect the number of reads mapping to this region. The dominant crRNA length is 66 nt, and each mature crRNA begins with the characteristic eight nucleotide handle at its 5’ end and ends with the remaining 22 nucleotides of the repeat. Both pathways produce the same mature crRNA, B, two types of icrRNA are generated. In ΔC × pTA232-telecrRNA, in addition to the 66-nucleotide-long crRNA, a shorter crRNA of 49 nt is also evident (Fig. 1C). RNAseq data from the longer icrRNA fraction (sizes 60–75 nucleotides) isolated from ΔC × pTA232-telecrRNA Haloferax cells (upper row, “long icrRNA”) and from the shorter icrRNA fraction (sizes 45–55 nucleotides) (lower row, “short icrRNA”) were mapped to the CRISPR C locus. Each icrRNA begins with the characteristic eight-nucleotide 5’ handle, followed by the spacer sequence. In contrast to the long crRNAs, the shorter crRNAs contain only a five-nucleotide long 3’ handle.

**TABLE 2**

Interference test with the icrRNA

Targeting efficiencies of the icRNAs expressed from the high copy and low copy icRNA plasmids were analyzed. The targeting efficiency of the icRNAs expressed from the high copy icRNA plasmid were investigated in strain ΔC and Δcas6b. A successful interference reaction reduces the transformation rate by at least a factor of 0.01, demonstrating a high targeting efficiency of the icrRNA (21).

<table>
<thead>
<tr>
<th>Position changed</th>
<th>schematic drawing of icrRNA</th>
<th>reduction in transformation rate by factor</th>
</tr>
</thead>
<tbody>
<tr>
<td>wild type</td>
<td>2</td>
<td>0.001</td>
</tr>
<tr>
<td>nucleotide –8 mutated: G → A (variant 4)</td>
<td>no reduction, no interference</td>
<td></td>
</tr>
<tr>
<td>nucleotide –8 mutated: G → U (variant 5)</td>
<td>0.005</td>
<td></td>
</tr>
<tr>
<td>nucleotide –8 mutated: G → C (variant 6)</td>
<td>0.006</td>
<td></td>
</tr>
<tr>
<td>nucleotide –1 mutated: C → U (variant 8)</td>
<td>0.0009</td>
<td></td>
</tr>
<tr>
<td>nucleotide –1 mutated: C → G (variant 9)</td>
<td>0.006</td>
<td></td>
</tr>
<tr>
<td>nucleotide –1 mutated: C → A (variant 10)</td>
<td>0.0003</td>
<td></td>
</tr>
<tr>
<td>3 nucleotide internal deletion (variant 7)</td>
<td>no reduction, no interference</td>
<td></td>
</tr>
<tr>
<td>nucleotide –8 deleted (variant 17)</td>
<td>0.0003</td>
<td></td>
</tr>
<tr>
<td>nucleotides –8 and –7 deleted (variant 18)</td>
<td>no reduction, no interference</td>
<td></td>
</tr>
<tr>
<td>5’ handle completely deleted (variant 3)</td>
<td>no reduction, no interference</td>
<td></td>
</tr>
</tbody>
</table>

ΔC: all Cascades in Δcas6b contain the icrRNA directed against the invader plasmid, whereas in ΔC only a percentage of the Cascade complexes are loaded with an icrRNA because the crRNAs from CRISPR locus P1 and P2 are also present. Taken together, the Cas6b protein is not required for the interference reaction when the icrRNA is present.

**Essential Features of the crRNA 5’ Handle—**Because the icrRNA was proven to be identical to the “naturally” expressed crRNA and to be fully active in interference, we generated different versions of the icrRNA to analyze the essential features of a crRNA for the interference reaction. To identify the important nucleotides of the 5’ handle, we generated 10 different variants and analyzed them for activity in the interference reaction (Table 3). All variants were transformed into strain Δcas6b that was subsequently challenged with the invader plasmid. First, we mutated the first nucleotide of the crRNA (which is a G) to a A, U, or C (variants 4 – 6). Mutation of the first nucleotide (position –8) results in icRNAs that are as effective in interference as the wild type icrRNA (Table 3). This is in agreement with the in vivo situation in Haloferax, where the crRNAs are generated from three different CRISPR loci, each of which have a different nucleotide at position –8 of the 5’ handle (Fig. 1A). Second, the –1 nucleotide was mutated from C to U and G and A (variants 8 – 10). This nucleotide has been shown in
E. coli (type I-E) to be derived from the invader (47–49). In *Haloferax*, the nucleotide –1 is a C and thus also identical to the last nucleotide of the PAM used in this study (TTC). It therefore has the potential to base pair with the invader (Fig. 3A). Mutation of this nucleotide to a U does not interfere with the defense activity. The U at this position could still base pair with the complementary PAM sequence in the invader (U-G base pair) (Fig. 3B). Mutation of the –1 nucleotide to a G, however, abolishes the defense activity, and this nucleotide could not base pair any longer with the complementary PAM sequence (GXG) (Fig. 3C). Surprisingly, mutation of this nucleotide from C to an A does not interfere with the defense activity, although an A at this position is not able to base pair with the complementary PAM sequence in the invader (GX A) (Fig. 3D).

Because the nature of the first crRNA nucleotide is not important, we next deleted this nucleotide, generating an icrRNA that is still active in interference. Deletion of the first two nucleotides results however in an icrRNA inactive in interference. A deletion of three nucleotides in the 5′ handle (positions –6 to –4) (variant 7) is not tolerated. The complete removal of the 5′ handle (variant 3) results in a crRNA that cannot trigger the interference reaction anymore.

Taken together, mutations in the 5′ handle are tolerated at the first nucleotide (position –8) and to some extent at position –1. Only the deletion of the first nucleotide of the 5′ handle is tolerated, and all other deletions result in inactive icRNAs.

**Essential Features of the crRNA 3′ Handle**—The crRNA 3′ handle in *Haloferax* has the potential to form a short stem loop structure at the very 3′ end (Fig. 1A). To determine whether parts of this stem loop are required and to define the essential features of the 3′ handle, we constructed 13 icrRNA variants with mutations in the 3′ handle and analyzed their activity in interference (Table 4). We mutated a nucleotide in the loop of the potential stem loop structure (G to C or U) (variants 11 and 12). These variants were both still active in triggering the interference reaction. The removal of four nucleotides of the 3′ handle in variant 1 (positions 8–11 in the 3′ handle) also did not interfere with the interference reaction. Likewise, the removal of 11 nucleotides in variant 2 (positions 1–11) did not reduce the interference. The nature of the 3′ handle differs from CRISPR system to CRISPR system. In *Haloferax* wild type cells, two types of crRNAs are observed having a 3′ handle of ~22 nucleotides and ~5 nucleotides (22). A similar observation was made with the icrRNA, because a long and a short icrRNA can be detected (Figs. 1C and 2B) that contains a 22-nucleotide and a 5 nucleotide 3′ handle (Fig. 2B). To investigate how many nucleotides can be removed from the 3′ handle, we designed several 3′ handle deletion variants. The five terminal nucleotides were

<table>
<thead>
<tr>
<th>position changed</th>
<th>schematic drawing of icrRNA variants</th>
<th>reduction in transformation rate by factor</th>
</tr>
</thead>
<tbody>
<tr>
<td>wild type</td>
<td><img src="image" alt="Schematic drawing of wild type icrRNA" /></td>
<td>0.001</td>
</tr>
<tr>
<td>mutation in loop: G → U (variant 11)</td>
<td><img src="image" alt="Schematic drawing of variant 11 icrRNA" /></td>
<td>0.003</td>
</tr>
<tr>
<td>mutation in loop: G → C (variant 12)</td>
<td><img src="image" alt="Schematic drawing of variant 12 icrRNA" /></td>
<td>0.005</td>
</tr>
<tr>
<td>deletion of nucleotides 8–11 (variant 1)</td>
<td><img src="image" alt="Schematic drawing of variant 1 icrRNA" /></td>
<td>0.008</td>
</tr>
<tr>
<td>deletion of nucleotides 1–11 (variant 2)</td>
<td><img src="image" alt="Schematic drawing of variant 2 icrRNA" /></td>
<td>0.002</td>
</tr>
<tr>
<td>5 nucleotide deletion from 3′ end (variant 13)</td>
<td><img src="image" alt="Schematic drawing of variant 13 icrRNA" /></td>
<td>0.01</td>
</tr>
<tr>
<td>10 nucleotide deletion from 3′ end (variant 14)</td>
<td><img src="image" alt="Schematic drawing of variant 14 icrRNA" /></td>
<td>0.008</td>
</tr>
<tr>
<td>15 nucleotide deletion from 3′ end (variant 15)</td>
<td><img src="image" alt="Schematic drawing of variant 15 icrRNA" /></td>
<td>0.004</td>
</tr>
<tr>
<td>20 nucleotide deletion from 3′ end (variant 16)</td>
<td><img src="image" alt="Schematic drawing of variant 16 icrRNA" /></td>
<td>0.002</td>
</tr>
<tr>
<td>1 nucleotide from 3′ end (variant 20)</td>
<td><img src="image" alt="Schematic drawing of variant 20 icrRNA" /></td>
<td>0.001</td>
</tr>
<tr>
<td>nucleotide 1 mutated G→C (variant 21)</td>
<td><img src="image" alt="Schematic drawing of variant 21 icrRNA" /></td>
<td>0.002</td>
</tr>
<tr>
<td>nucleotide 1 mutated G→A (variant 22)</td>
<td><img src="image" alt="Schematic drawing of variant 22 icrRNA" /></td>
<td>0.005</td>
</tr>
<tr>
<td>nucleotide 1 mutated G→U (variant 23)</td>
<td><img src="image" alt="Schematic drawing of variant 23 icrRNA" /></td>
<td>0.008</td>
</tr>
<tr>
<td>no 3′ handle (variant 19)</td>
<td><img src="image" alt="Schematic drawing of variant 19 icrRNA" /></td>
<td>0.001</td>
</tr>
</tbody>
</table>
removed in variant 14, and the last 15 and 20 nucleotides were deleted in variants 15 and 16, respectively. The interference tests clearly show that all four deletions in the 3’ handle had no effect on the interference activity (Table 4). In variant 20, only one nucleotide of the 3’ handle remained, but still this crRNA was effective in triggering the interference reaction. This last nucleotide was mutated in variants 21–23 from a G to a C, A, or U. Again, all variants were still active. Even a complete removal of all 22 nucleotides (variant 19) did not interfere with the interference reaction. These results also suggest that the exact length of the complete crRNA is not important, because different lengths at the 3’ handle are tolerated.

**DISCUSSION**

We could successfully establish a Cas6b-independent crRNA maturation pathway in *Haloferax* cells. In this pathway, icrRNAs are excised from a precursor with the help of tRNA processing enzymes, resulting in small RNAs active in the interference reaction. The icrRNAs are identical to the natural crRNAs except for the nature of the end groups.

*Cas6b Is Only Required for crRNA Maturation in Type I-B—*

Using the independently generated crRNA, we could show that Cas6b is not required for any other reactions besides crRNA processing in the prokaryotic immune system I-B. As soon as the crRNA is generated without Cas6b, this protein is dispensable, because it is not required for the interference reaction. We previously showed that Cas6b copurifies with Cascade in *Haloferax* (22), and this observation might be due to the fact that the crRNA is incorporated into Cascade and that Cas6b is still bound to the crRNA thereby co-purifying with the FLAG-tagged Cas7. But although it copurifies with Cas5 and Cas7, it is not required to be part of the I-B Cascade for activity. Thus, the core part of the I-B Cascade seems to consist of Cas5, Cas7, and the crRNA. These results are confirmed by the observation that the 3’ handle can be completely removed. Thus, if the Cas6b protein is attached to Cascade via binding to the crRNA 3’ handle, this interaction is not essential.

**Essential Parts of the 5’ Handle—** Recent reports on the structure of the *E. coli* Cascade complex revealed that the first seven nucleotides of the crRNA 5’ handle form a hook that interacts with the Cas5, Cas7, and Cse1 proteins (the homologous protein in *Haloferax* would be the Cas8b protein) (42–44). Our data clearly show that in the *Haloferax* I-B system, the 5’ handle is also an important part of the crRNA. Only the first nucleotide of the 5’ handle can be mutated and deleted without loss of activity. This is in agreement with the *in vivo* situation where three different 5’ handles are generated (Fig. 1A). In the structural analyses reported for the I-E Cascade complex, the first nucleotide of the 5’ handle interacts with Cas5 and Cas7 (42–44). In the *Haloferax* system, this interaction does not seem to be crucial for the activity. However, all other deletions in the 5’ handle abolished interference activity as follows: deletions of the first two nucleotides, of three internal nucleotides, and of all 5’ handle nucleotides yield a nonfunctional crRNA, confirming the importance of the 5’ handle.

**Interaction of the crRNA 5’ Handle with the Protospacer Adjacent Motif—** The nature of the last nucleotide of the 5’ handle (position −1) seems to be important; mutation of this nucleotide from C to G results in loss of activity, and only nucleotides C, A, and U are tolerated at this position. In *E. coli*, it has been shown that the −1 crRNA nucleotide is identical to the last PAM nucleotide and is derived from the invader (47–51), and thus the crRNA could base pair with the invader at this position (Fig. 3). It is not known whether the crRNA 5’ handle nucleotide (position −1) stems from the invader in the *Haloferax* I-B system. But the −1 crRNA nucleotide and invader complementary PAM nucleotide (in PAMs TTC and CAC, two of the six PAMs recognized by *Haloferax*) also have the potential to base pair. This base pair might be important for recognizing the correct target DNA sequence. The observation that the −1 nucleotide mutant C to U works but C to G does not work would confirm this hypothesis. However, the result that the C to A mutation is still active in interference does not fit. In addition, the complementary nucleotide of the other four PAMs recognized in the *Haloferax* system (TAT, TAA, TAG, and ACT) cannot base pair with the crRNA. In the I-E and I-F *E. coli* system, it has been shown that the interaction between the −1 crRNA nucleotide and the last complementary PAM nucleotide is not essential for invader recognition (52, 53). The recent structural data for the I-E Cascade complex confirm this earlier observation showing that in this system the −1 nucleotide of the crRNA is displaced by the Cas5 protein preventing interaction with the invader PAM sequence. The same displacement of the −1 nucleotide might happen in the *Haloferax* I-B Cascade. Also, the loss of activity of the C→G mutant could be explained by failure of the G to interact properly with the Cas5 protein.

In the I-E system, Cse1 (the homologous protein in *Haloferax* is Cas8b) interacts with the PAM sequence, and target recognition occurs via identification of the PAM sequence by the Cse1 protein (18, 53–55). The same might be true for the *Haloferax* I-B system, but the Cas8b protein should be able to identify six different PAMs as follows: TTC, CAC, TAT, TAA, TAG, and ACT. Taken together, our results suggest that a G at position −1 cannot interact properly with the Cas5 protein and that the *Haloferax* Cas8b would have to recognize all six different PAMs.

**Essential Parts of the 3’ Handle—** Mutational analysis of the icrRNA showed that the 3’ handle of the crRNA is completely dispensable. The shortest icrRNA found *in vivo* by RNAseq contained a five-nucleotide-long 3’ handle. According to the data reported here, this shorter crRNA version with only 49 nucleotides should also be active, because even an icrRNA with no 3’ handle is still active. Previous isolation of icrRNAs from the *Haloferax* Cascade-like complex showed that the long and the short crRNA versions co-purify (22). It would be interesting to analyze whether only the short form is the active form and whether the long form has to be activated by 3’-processing to yield the short functional form. Currently, it is not known which enzyme(s) are catalyzing this further trimming of the crRNA 3’ end. As soon as this enzyme is identified, we could generate a strain that has the gene for the enzyme deleted and analyze whether the icrRNA with a long unprocessed 3’ handle is active.

A shortening of the crRNA 3’ handle has also been reported for the type I-B system of *Methanococcus maripaludis* and
Clostridium thermocellum (56). Thus, it seems that in contrast to the I-A, I-E, and I-F systems, crRNAs of the I-B system are subjected to an additional 3' trimming, as reported for the crRNAs in type III systems (14, 15).

Nature of the crRNA End Group Is Not Important—The pre-crRNA is generated by the tRNA-processing enzymes to exactly the same product as the pre-crRNA generation by Cas6. The only difference between the natural crRNA and the icrRNA is the nature of the 5' handle and a 34-nucleotide-long spacer.

Minimal crRNA—Previously published data concerning the requirements for the spacer-protospacer interactions in the Haloferax I-B system showed that a 34-nucleotide-long spacer-protospacer interaction between crRNA and invader was sufficient (16). According to these published data and the results reported here, the minimal crRNA for the Haloferax type I-B system contains a 7-nucleotide-long 5' handle, a 34-nucleotide-long spacer, and no 3' handle (Fig. 4). Altogether, this crRNA would be 41 nucleotides long.

Acknowledgments—We are grateful to Charles J. Daniels for providing us with the synthetic promoter and terminator sequences for Haloferax ahead of publication. We thank Eli Bruckbauer for expert technical assistance and Jutta Brendel for the kind help. We are grateful to all members of the FOR Unit 1680 for helpful discussions.

REFERENCES
Prokaryotic Immune Defense with an icrRNA and without Cas6


30. Stoll, B. (2013) Biology II, Ulm University, Ulm, Germany


