Landes Bioscience www.landesbioscience.com

Supplemental Material to:

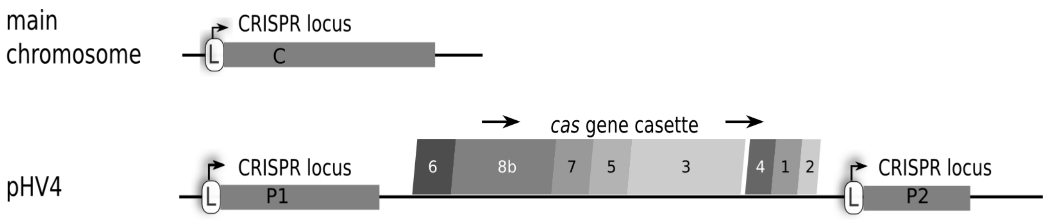
Lisa-Katharina Maier, Sita Lange, Britta Stoll, Karina Haas, Susan Fischer, Eike Fischer, Elke Duchardt-Ferner, Jens Wöhnert, Rolf Backofen and Anita Marchfelder

Essential requirements for the detection and degradation of invaders by the *Haloferax volcanii* CRISPR/Cas system I-B

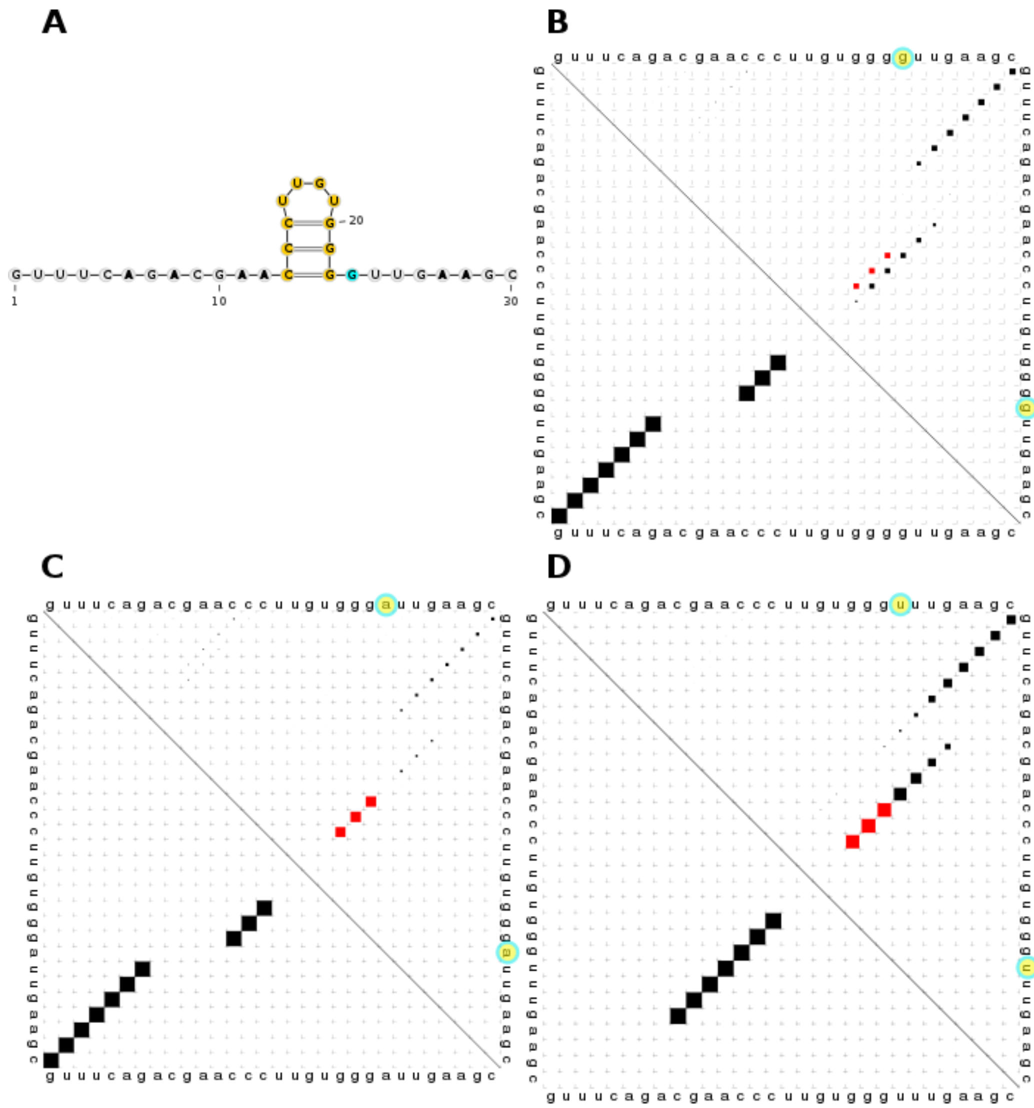
2013; 10(5)
http://dx.doi.org/10.4161/rna.24282

www.landesbioscience.com/journals/rnabiology/article/24282/

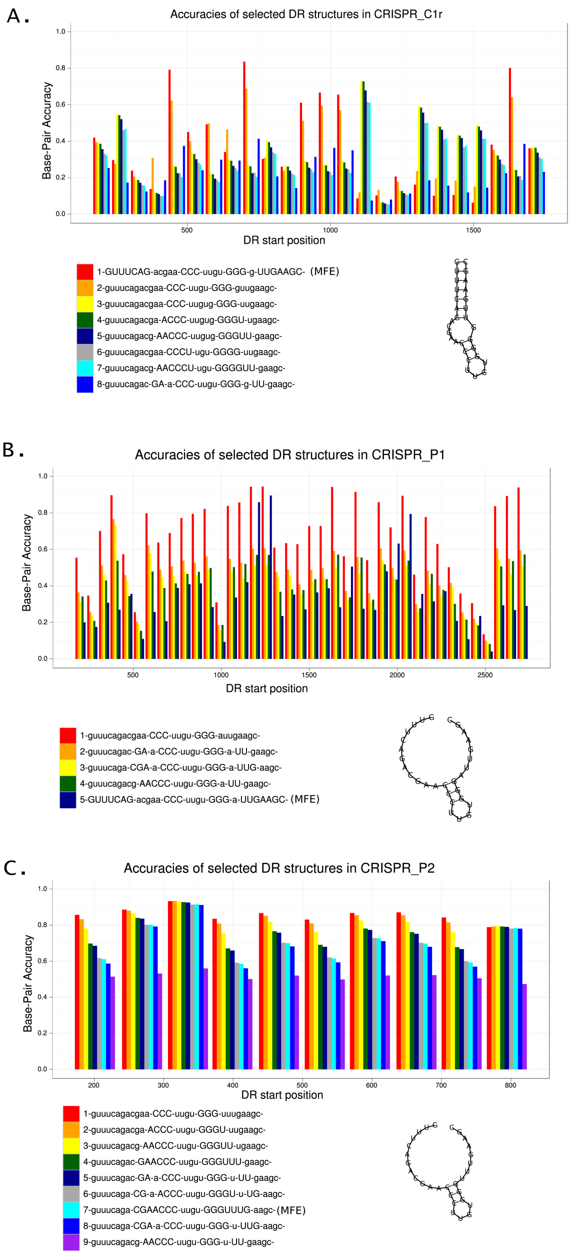
Supplementary figures and table



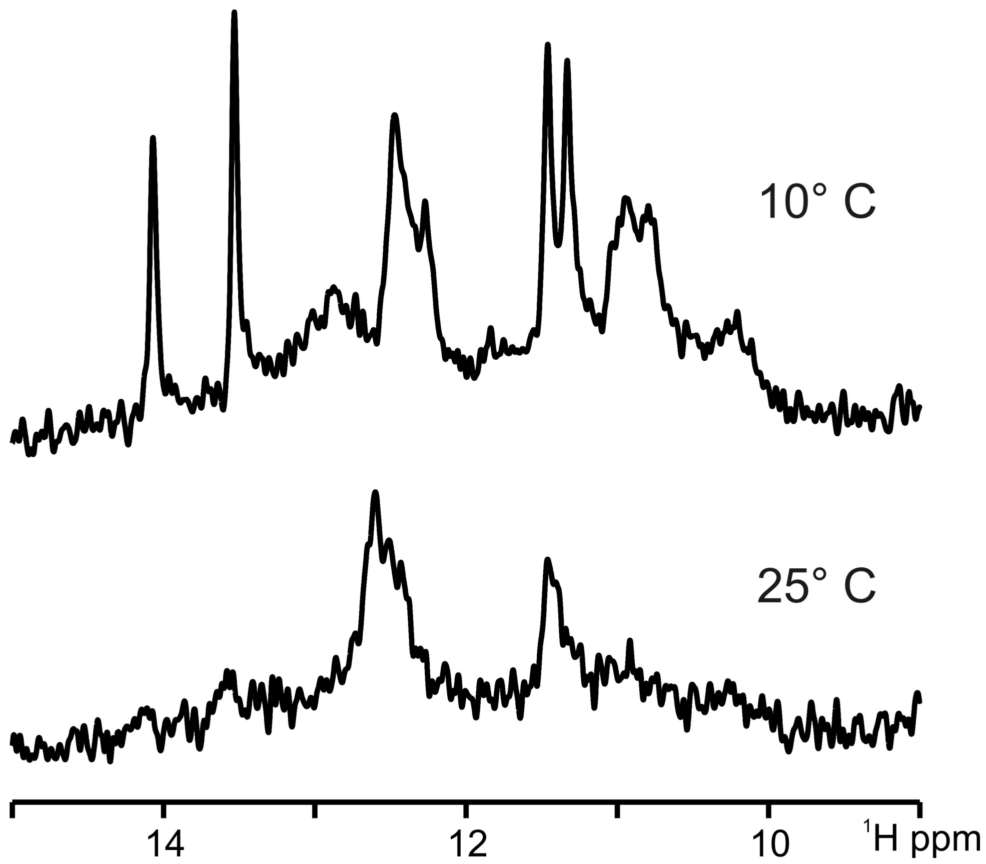
Supplementary Figure 1. The *Haloferax volcanii* CRISPR/Cas system. One CRISPR RNA locus (Locus C) is encoded on the main chromosome and contains 25 repeats and 24 spacers. The leader region (L) at the 5´ end of the CRISPR locus contains the promoter (black arrow). Two CRISPR RNA loci (P1 and P2) are flanking the *cas* gene cluster on the mini chromosome pHV4. CRISPR locus P1 contains 17 repeats and 16 spacers, and locus P2 contains 12 repeats and 11 spacers. The *cas* gene cluster codes for the proteins Cas1, Cas2, Cas3, Cas4, Cas5, Cas6, Cas7 and Cas8b, the latter being the signature protein for subtype I-B.



Supplementary Figure 2. The structure motif containing three C-G base pairs forms in all CRISPR loci when considering the entire repeat-spacer context. A. The conserved stem-loop is stable across all repeat loci (highlighted as red dots in B-D). **B.-D.** Dotplots to determine the influence of the array sequence on the individual repeat instances are shown. A dotplot is a matrix where the dot (i,j) corresponds to the base pair of the nucleotides at position i and j in the sequence and its size is relative to the base pair probability; the larger the dot, the more probable the base pair is. The lower triangle depicts a dot for the minimum free energy structure as predicted by RNAfold. The upper triangle depicts the average base pair probabilities for the repeat sequence in each instance in its respective repeat-spacer array. The array was folded with RNAplfold. This means that the upper triangle determines the average stability of possible base pairs across the entire array. This is in contrast to the traditional approach of just folding the repeat sequence on its own. However, the surrounding sequence influences the stability of each individual repeat instance in the array.



Supplementary Figure 3. Repeat folding is influenced by the neighbouring spacer sequences. For all repeats of all three spacer loci the potential of the repeat sequence to fold into a certain structure is shown. Depending on the adjacent spacer sequences the repeat is predicted to fold into the given structures with different stabilities (measured as the average base pair probability for a given structure, also called structure accuracy). Nucleotides involved in base pairing are shown in upper case letters. **A.** Folding of the repeats from CRISPR locus C. **B.** Folding of the repeats from CRISPR locus P1. **C.** Folding of the repeats from CRISPR locus P2. In each case the secondary structure of the structure coded in red is shown in the lower right corner.



Supplementary Figure 4. NMR analysis of the P1 repeat. Imino proton region of 1D-1H-NMR-spectra at 10°C (top) and 25°C (bottom). At 25°C only two very broad imino proton signals barely above the noise level are visible indicating the complete absence of stable base pairing interactions. At 10°C additional signals become observable. However, the large line widths for some signals as well as the strong inhomogeneity in imino proton line-widths argue against the presence of a structurally well-defined fold of the repeat RNA.

Supplementary Table 1. Strains, plasmids and primers used in this study. The PAM sequences are underlined in the primer sequences, mutated nucleotides are shown in bold.

Strains	**Genotype**	**Source/Reference**
DH5α	F- φ80*lac*ZΔM15 Δ(*lac*ZYA-*arg*F) U169 *rec*A1 *end*A1 *hsd*R17 (rk-, mk+) *gal*- *pho*A *sup*E44 λ- *thi*-1 *gyr*A96 *rel*A1	(invitrogen)
GM121	F− *dam-3 dcm-6 ara-14 fhuA31 galK2 galT22 hdsR3 lacY1 leu-6 thi-1 thr-1 tsx-78*	(42)
H119	Δ*pyrE2* ΔtrpA Δ*leu*B	(26)
H26	Δ*pyrE2*	(26)
Plasmids	**Relevant properties**	**Source/Reference**
pTA409	Shuttle vector with *pyrE2* marker and pHV1 replication origin	(24)
pTA409-PAM3	Spacer P1.1. downstream of PAM3 (TTC)	(11)
pTA409-PAM9	Spacer P1.1. downstream of PAM9 (ACT)	(11)
pTA409-PAM3-P1.X/P2.X/C.X	Spacers number X of locus P1, P2 or C preceded by PAM3(TTC)	This study
pTA409-PAM9-P1.X/P2.X/C.X	Spacers number X of locus P1, P2 or C preceded by PAM9(ACT)	This study
pTA352	Shuttle vector with *leu*B marker and pHV1 replication origin	(23)
pTA352-PAM3	Spacer P1.1. downstream of PAM3 (TTC)	This study
pTA352-PAM9	Spacer P1.1. downstream of PAM9 (ACT)	This study
pTA232	Shuttle vector with *leu*B marker and pHV2 replication origin	(26)
pTA232-PAM3	Spacer P1.1. downstream of PAM3 (TTC)	This study
pTA232-PAM9	Spacer P1.1. downstream of PAM9 (ACT)	This study
pTA409-SEED1	Spacer P1.1 mutated at position 1 downstream of PAM9 (ACT)	This study
pTA409-SEED2-25	Spacer P1.1 mutated at positions indicated in Figure 3 downstream of PAM9 (ACT)	This study
Primers	**Sequence**	**Source/Reference**
VPAM9	TACTGCAGGCATCTCGACCGGCGACCTCC	(11)
VPAM9mutP1	TACT**A**CAGGCATCTCGACCGGCGACCTCC	This study
VPAM9mut1(P2)	TACTG**A**AGGCATCTCGACCGGCGACCTCC	This study
VPAM9mutP3	TACTGC**C**GGCATCTCGACCGGCGACCTCC	This study
VPAM9mutP4	TACTGCA**T**GCATCTCGACCGGCGACCTCC	This study
VPAM9mutP5	TACTGCAG**A**CATCTCGACCGGCGACCTCC	This study
VPAM9mutP6	TACTGCAGG**T**ATCTCGACCGGCGACCTCC	This study
VPAM9mutP7	TACTGCAGGC**G**TCTCGACCGGCGACCTCC	This study
VPAM9mutP8	TACTGCAGGCA**G**CTCGACCGGCGACCTCC	This study
VPAM9mutP9	TACTGCAGGCAT**A**TCGACCGGCGACCTCC	This study
VPAM9mutP10	TACTGCAGGCATC**G**CGACCGGCGACCTCC	This study
VPAM9mutP11	TACTGCAGGCATCT**A**GACCGGCGACCTCC	This study
VPAM9mutP12	TACTGCAGGCATCTC**T**ACCGGCGACCTCC	This study
VPAM9mutP13	TACTGCAGGCATCTCG**G**CCGGCGACCTCC	This study
VPAM9mutP14	TACTGCAGGCATCTCGA**T**CGGCGACCTCC	This study

VPAM9mutP15	TACTGCAGGCATCTCGAC**A**GGCGACCTCC	This study
VPAM9mutP16	TACTGCAGGCATCTCGACC**A**GCGACCTCC	This study
VPAM9mut3(P17)	TACTGCAGGCATCTCGACCG**A**CGACCTCC	This study
VPAM9mutP18	TACTGCAGGCATCTCGACCGG**A**GACCTCC	This study
VPAM9del17	TACTGCAGGCATCTCGACCGCGACCTCC	This study
VPAM9P12,17	TACTGCAGGCATCTC**T**ACCG**A**CGACCTCC	This study
Spacer1rev	CAAAGTGTTCCGGGAGGTCGCCGGTC	(11)
HmutP12	ATCAAAGTGTTCCGGGAGGTCGCCGGT**A**	This study
HmutP13	ATCAAAGTGTTCCGGGAGGTCGCCGG**C**C	This study
HmutP14	ATCAAAGTGTTCCGGGAGGTCGCCG**A**TC	This study
HmutP15	ATCAAAGTGTTCCGGGAGGTCGCC**T**GTC	This study
HmutP16	ATCAAAGTGTTCCGGGAGGTCGC**T**GGTC	This study
HPAM9mut3(P17)	ATCAAAGTGTTCCGGGAGGTCG**T**CGGTC	This study
HPAM9mut2(P36)	TACTC**G**AAGTGTTCCGGGAGGTCGCCGGTC	This study
HmutP18	ATCAAAGTGTTCCGGGAGGTC**T**CCGGTC	This study
Hdel17	ATCAAAGTGTTCCGGGAGGTCGCGGTC	This study
HmutP12,17,36	TAC**G**AAGTGTTCCGGGAGGTCG**T**CGGT**A**	This study
HmutP12,17,26,36	ATC**G**AAGTGTTCC**T**GGAGGTCG**T**CGGT**A**	This study
HmutP17,36	ATC**G**AAGTGTTCCGGGAGGTCG**T**CGGTC	This study
HmutP31-33	ATCAAA**ACA**TTCCGGGAGGTCGCCGGTC	This study
HmutP31-37	AT**TGCGACA**TTCCGGGAGGTCGCCGGTC	This study
HmutP35-37	AT**TGC**AGTGTTCCGGGAGGTCGCCGGTC	This study
V PAM3 P1-Mitte	TTTCGGGAGTGTATGCGATATCCCTTAAAAC	This study
V PAM9 P1-Mitte	TACTGGGAGTGTATGCGATATCCCTTAAAAC	This study
H PAM9 P1-Mitte	AAGTAGGGGGTTTTAAGGGATATCG	This study
V PAM3 P1- Ende	TTTCGTCGCTCGCGGGATCGACATCATCGC	This study
V PAM9 P1- Ende	TACTGTCGCTCGCGGGATCGACATCATCGC	This study
H PAM9 P1- Ende	CATCGCCTTGAGCGCGATGATGTCGATC	This study
V PAM3 P2-Sp1	TTTCTCGGGGTGGACGTTCTGCCGGGCATAG	This study
V PAM9 P2- SP1	TACTTCGGGGTGGACGTTCTGCCGGGCATAG	This study
H PAM9 P2-SP1	CAATATATTGCCTATGCCCGGCAG	This study
V PAM3 P2-Sp6	TTTCTGCTAGTCGGACCGAGGGAGGAAGC	This study
V PAM9 P2-SP6	TACTTGCTAGTCGGACCGAGGGAGGAAGC	This study
H PAM9 P2-SP6	GATCCGAGAGGCTTCCTCCCTCG	This study
VPAM3 P2Sp8	TTTCTGTGGGGAACAAGCAATAGAGG	This study
VPAM9 P2Sp8	TACTTGTGGGGAACAAGCAATAGAGG	This study
H P2Sp8	CATCCTTCAAAGTCCTCTATTGCTTGT	This study
V PAM3 P2-SP10	TTTCGGCGACAACGACGGCAACCTCGAAG	This study
V PAM9 P2-SP10	TACTGGCGACAACGACGGCAACCTCGAAG	This study
H PAM9 P2-SP10	AGGCGGCGCAGACTTCGAGGTTGC	This study
V PAM3 P2-SP10	TTTCGGCGACAACGACGGCAACCTCGAAG	This study
V PAM9 P2-SP10	TACTGGCGACAACGACGGCAACCTCGAAG	This study
H PAM9 P2-SP10	AGGCGGCGCAGACTTCGAGGTTGC	This study
V PAM3 C-Sp1	TTTCGAGGAAAGGTCCGAAGATGCCCTCG	This study
V PAM9 C- SP1 NEU	TACTGAGGAAAGGTCCGAAGATGCCCTCG	This study
H PAM9 C- SP1 NEU	ATTTTCTGAGATTCGAGGGCATCTTC	This study
VPAM3 CSp9	TTTCGCGGAGGTCGCTGACAGCAACG	This study
VPAM9 CSp9	TACTGCGGAGGTCGCTGACAGCAACG	This study
H CSp9	CGCTCGATACCGTCGTTGCTGTCAGCGA	This study

V PAM3 C-Sp10	TTTCAAAATCCTCGATTACTACGAGATCCC	This study
VPAM9 SP10C3	TACTAAAATCCTCGATTACTACGAGATCCC	This study
H-SP10C3	ATAATCTTCCGTGGGATCTCGTAGTAATC	This study
V PAM3 C-SP24 NEU	TTTCTCGACTAACGTGGTCTCGTTCGG	This study
V PAM9 C-SP24 NEU	TACTTCGACTAACGTGGTCTCGTTCGG	This study
H PAM9 C-SP24 NEU	ATATCCACCGTGCCGCCGAACGAGACCA	This study