SUPPLEMENTARY DATA

Modulating the Cascade architecture of a minimal Type I-F CRISPR-Cas system

Daniel Gleditzsch[1], Hanna Müller-Esparza[1], Patrick Pausch[2,3], Kundan Sharma[4], Srivatsa Dwarakanath[1], Henning Urlaub[4,5], Gert Bange[2,3] and Lennart Randau[1,2,*]

Supplementary Tables

Supplementary Table S1. List of constructed plasmids and investigated crRNA sequences

Plasmid	Description
pCDFDuet-1	T7 RNAP polymerase based expression vector, Spec^R (Novagen)
pRSFDuet-1	T7 RNAP polymerase based expression vector, Kan^R (Novagen)
pUC19	Cloning vector, Amp^R (New England Biolabs)
pCas1	pRSFDuet-1 + *cas7fv*, *cas5fv* and *cas6f* with an N-terminal His-tag on *cas7fv*
pCas2	pRSFDuet-1 + *cas7fv*, *cas5fv* and *cas6f* with a C-terminal His-tag on *cas6f*
pCas3	pRSFDuet-1 + *cas7fv* and *cas6f* with an N-terminal His-tag on *cas7fv*
pCas4	pRSFDuet-1 + *cas3*, *cas7fv*, *cas5fv* and *cas6f* with an N-terminal His-tag on *cas3*
pCas5	pRSFDuet-1+ *cas1, cas3*, *cas7fv*, *cas5fv* and *cas6f* with an N-terminal His-tag on *cas1*
pCas6	pRSFDuet-1+ *cas1, cas3*, *cas7fv*, *cas5fv* and *cas6f*
pCas7	pCas3 with a C466G, A467C, A470C mutation in *cas3*
pCRISPR-wt	pUC19 + T7 RNAP promoter and the repeat-spacer4-repeat sequence of *S.putrefaciens* CN-32 Spacer 4: 5′-TATCGCCCAGCAAGACGCGCAAACCTATAACC-3′
pCRISPR+18	pCRISPR-wt with spacer4 extended by 18 random nt: 5′-CACTCAATCCCGCTAATG-spacer4-3′
pCRISPR-18	pCRISPR-wt with spacer4 shortened by 18 nt
pCRISPR+15	pCRISPR-wt with spacer4 extended by 15 random nt: 5′-CACTCAACGCTAATG-spacer4-3′
pCRISPR-15	pCRISPR-wt with spacer4 shortened by 15 nt
pCRISPRλ	pCDFDuet-1+ T7 RNAP promoter and a repeat-spacer-repeat array. The repeat sequence was obtained from *S. putrefaciens* CN-32. The 32 nt spacer is complementary to a phage lambda gene E sequence and flanked by a GG PAM at the 3´end of the target strand. : 5′-GGCGGCACGGAGTGGAGCAAGCGTGACAAGTC-3′
pCRISPRλGA	pCDFDuet-1+ T7 RNAP promoter and a repeat-spacer-repeat array. The repeat sequence was obtained from *S. putrefaciens* CN-32.The 32 nt spacer is complementary to phage lambda gene E sequence and flanked by a GA PAM at the 3´end of the target strand: 5′-GCATCATCATGCAGAACATGCGTGACGAAGAG-3′
pCRISPR λ+1	pCRISPR λ with a spacer sequence elongated by 1 nt complementary to phage lambda gene E sequence.
pCRISPR λ+6	pCRISPR λ with a spacer sequence elongated by 6 nt complementary to phage lambda gene E sequence.
pCRISPR λ+9	pCRISPR λ with a spacer sequence elongated by 9 nt complementary to phage lambda gene E sequence.
pCRISPR λ+12	pCRISPR λ with a spacer sequence elongated by 12 nt complementary to phage lambda gene E sequence.
pCRISPR λ+18	pCRISPR λ with a spacer sequence elongated by 18 nt complementary to phage lambda gene E sequence.
pCRISPR λ-6	pCRISPR λ with a spacer sequence shortened by 6 nt
pCRISPR λ-18	pCRISPR λ with a spacer sequence shortened by 18 nt
pCRISPR NT	pCRISPR λ with a spacer sequence without a match in the phage lambda genome: 5′-GCATCATCATGCAGAACATGCGTGACGAAGAG-3′
Sp4-GG-target	Oligonucleotide used for EMSAs containing spacer 4 target sequence with a GG-PAM 5′-AAGCTTGAGGGCCCAAGCCGTTATGCTAGGGTTATAGGTTTGCGCGTCTTGCTGGGCGATAGG ACTCCCTATAGTGAGTCGTATTAGGATCC-3′
Sp1-GG-target	Oligonucleotide used for EMSAs containing spacer 1 target sequence with a GG-PAM 5′-AAGCTTGAGGGCCCAAGCCGTTATGCTAGCAATGTGGTCGCGCAATTTATGATTTGGTTGAGG ACTCCCTATAGTGAGTCGTATTAGGATCC-3′

Supplementary Table S2
Small angle X-ray scattering (SAXS) statistics

Data collection	**S Cascade**	**WT Cascade**	**L Cascade**
Detector	Pilatus 1M	Pilatus 1M	Pilatus 1M
Beam size at sample (µm µm)	700 x 700	700 x 700	700 x 700
Wavelength (Å)	0.9919	0.9919	0.9919
q range (nm^{-1})	0.025 - 5	0.025 - 5	0.025 - 5
Exposure time (s)	10 (10 x 1)	10 (10 x 1)	10 (10 x 1)
Concentration range (mg ml^{-1})	3 - 10	2.5 - 20	10 - 25
Temperature (°C)	4	4	4
Structural parameters			
Sample concentration (mg ml^{-1})	10	20	25
I(0) (cm^{-1}) from P(r)	95.54 +/- 0.0367	160.4 +/- 0.0622	86.72 +/- 0.04148
R_g (nm) from P(r)	4.173 +/- 0.002564	5.444 +/- 0.002526	6.524 +/- 0.003503
I(0) (cm^{-1}) from Guinier	95.53	160.4	86.72
R_g (nm) from Guinier	4.167	5.435	6.517
D_{max} (nm)	14.19	18.14	21.66
Porod volume estimate (nm^3)	207.82	360.28	519.09
***ab initio* modeling results**			
Normalized spatial distribution (NSD) and NSD variation	0.818 (0.044)	0.822 (0.039)	0.865 (0.063)
Final χ^2 against raw data	1.07	2.180	4.538
Software employed			
Primary data reduction	ESRF BM29 online/ Primus	ESRF BM29 online/ Primus	ESRF BM29 online/ Primus
Data processing	Primus	Primus	Primus
3D graphics representation	Chimera	Chimera	Chimera

Primus (44)(ATSAS software package) was used for data analysis. 20 *ab initio* shape restoration models have been calculated with Dammif (implemented in Primus) in slow mode to generate averaged and filtered models.

Supplementary Figures

Supplementary Figure S1

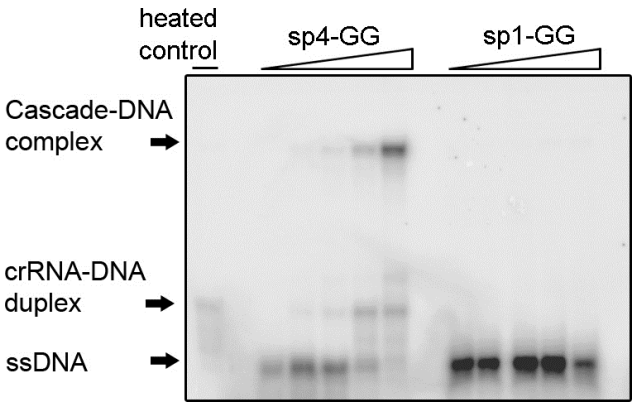


Fig. S1. Recombinant Type I-Fv Cascade binds ssDNA

5'-radioactively labeled target DNA strands (~0.4 pmoles) were incubated with increasing amounts (0, 2, 4, 20, 60 nM) of recombinant *S. putrefaciens* Type I-Fv Cascade containing a crRNA with spacer 4 of the *S. putrefaciens* CRISPR array. Binding was observed in electrophoretic mobility shift assays for a target DNA that was complementary to the Cascade-bound crRNA 4 (sp4-GG). A non-complementary sequence containing the target of spacer 1 and a "GG"-PAM (sp1-GG) was not bound. A control was included that contained 60 nM of Cascade after heat-incubation (10 min, 95°C) leading to Cas protein denaturation and crRNA release. This control identified a first smaller shift as a potential crRNA-ssDNA target duplex.

Supplementary Figure S2

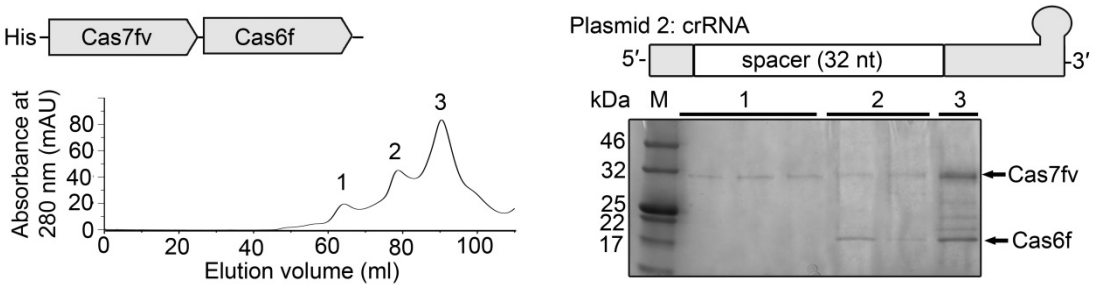


Fig. S2. Cascade is not produced in the absence of Cas5fv

(Top) Schematic overview of the investigated recombinant Cas proteins (pCas3) and crRNA. SDS-PAGE (right) was used to analyze the protein and RNA content corresponding to the peaks indicated in the gel-elution chromatogram (left).

Supplementary Figure S3

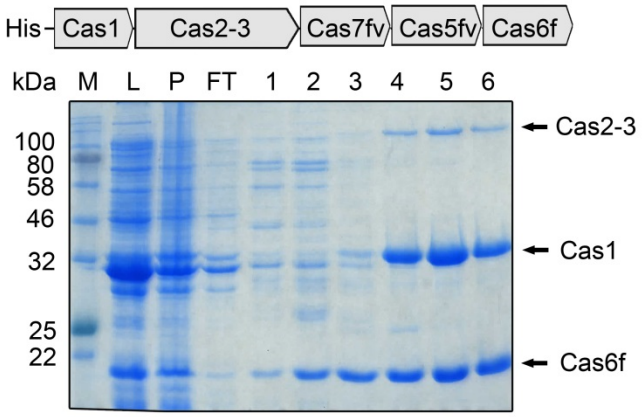
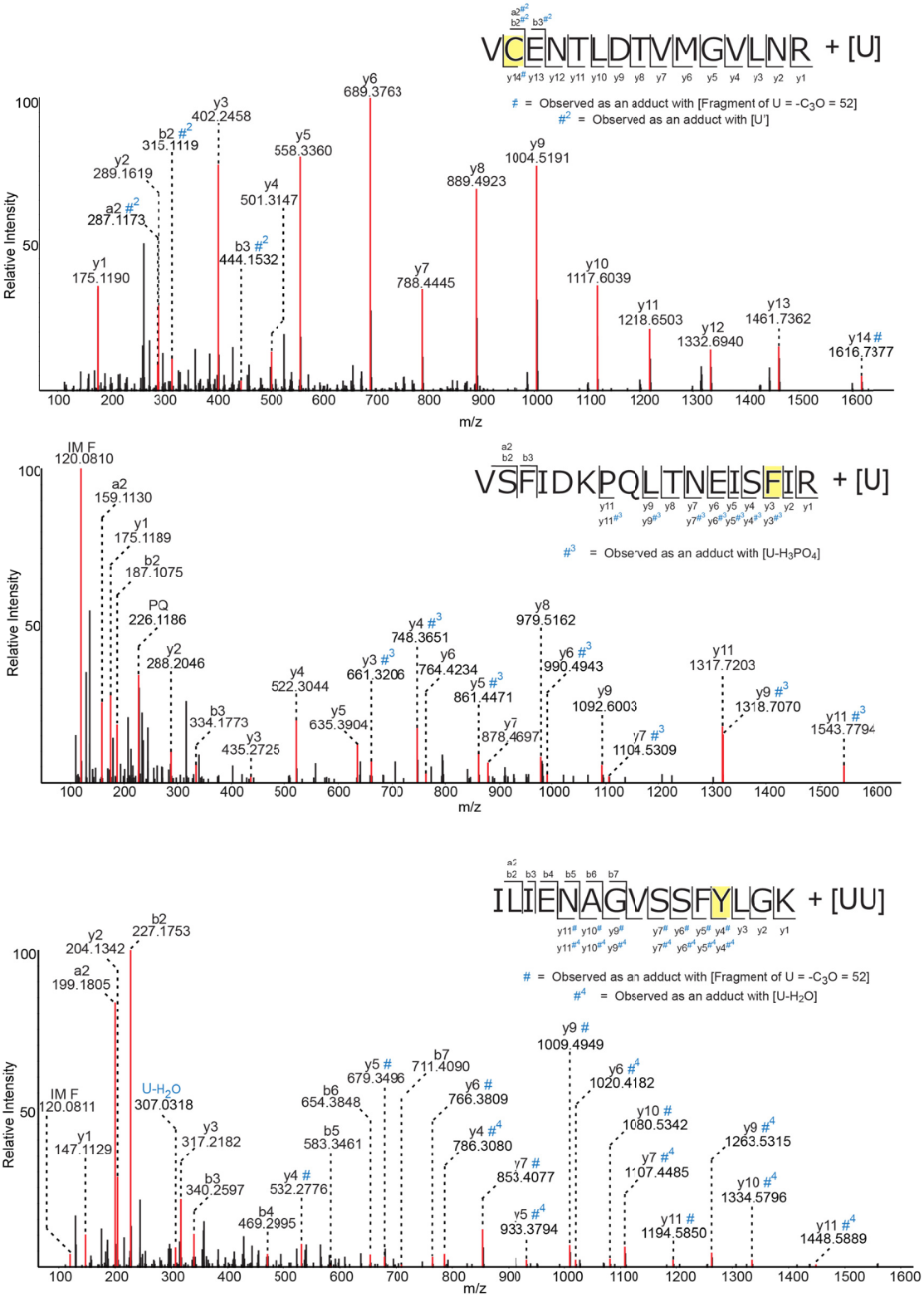


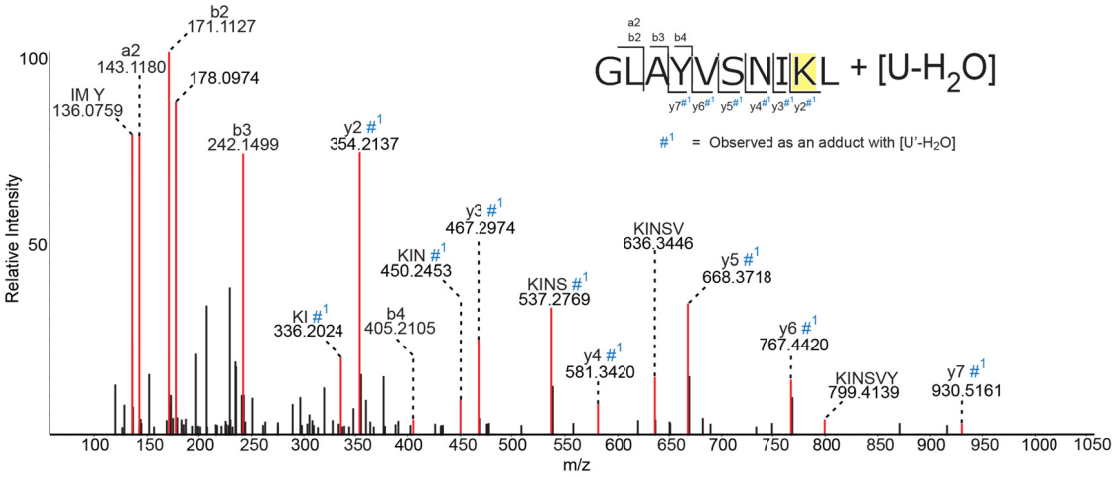
Fig. S3. Cas1 interacts with Cas3 and Cas6f but not with the Cascade backbone

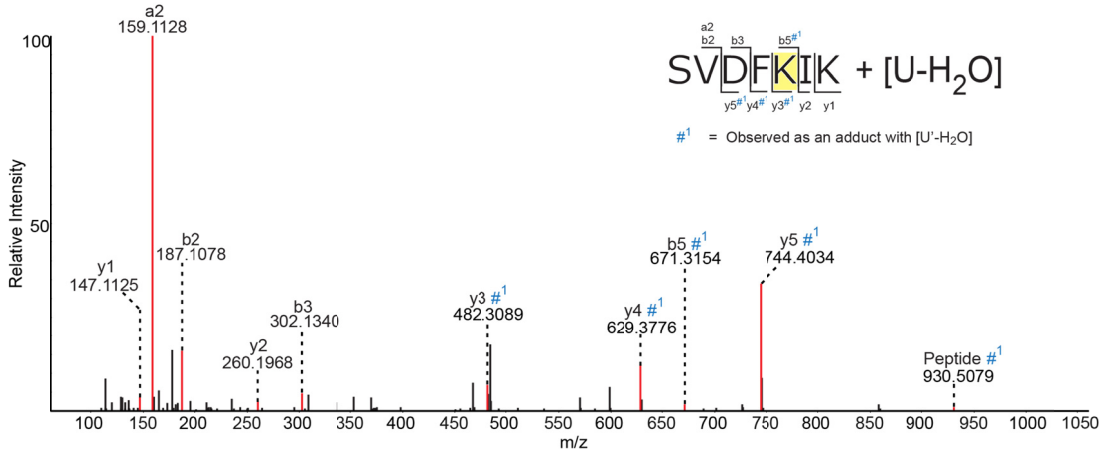
(Top) Schematic overview of the investigated recombinant Cas proteins (pCas5) and crRNA. SDS-PAGE (bottom) was used to analyze the protein content of the Ni-NTA purification fractions (Marker, M, Lysate, L, Pellet, P, Flow-through, FT).

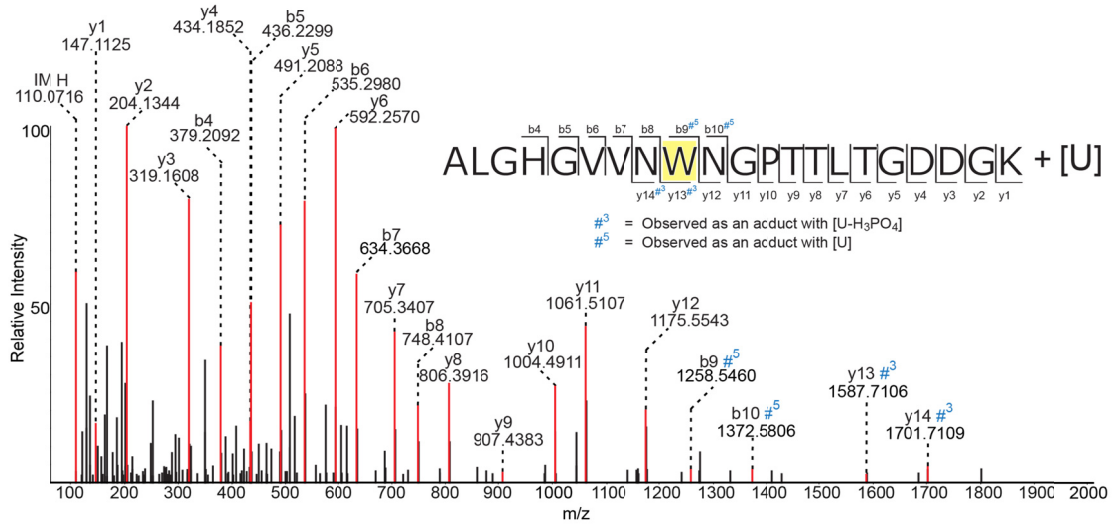
Supplementary Figure S4



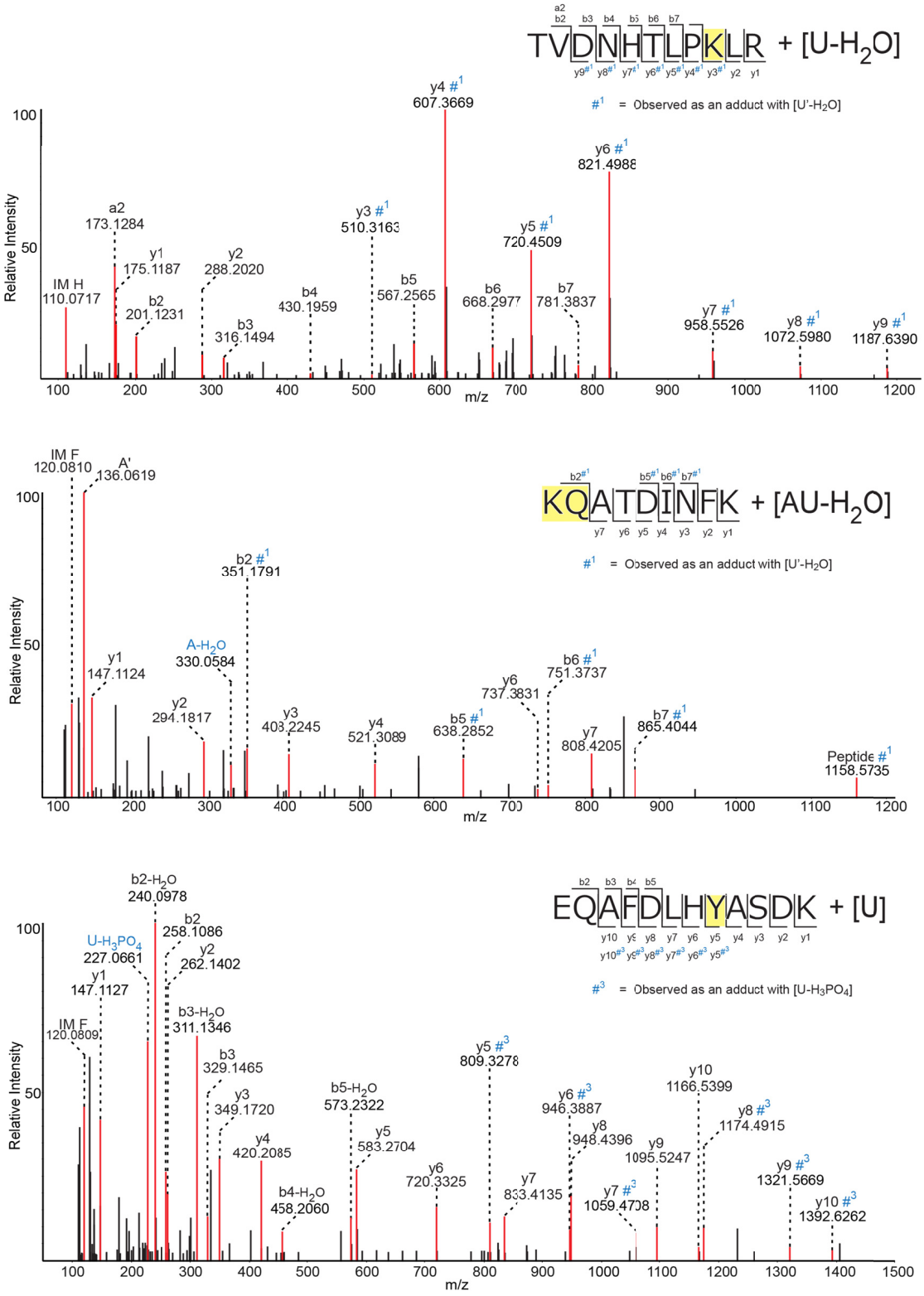
Supplementary Figure S5



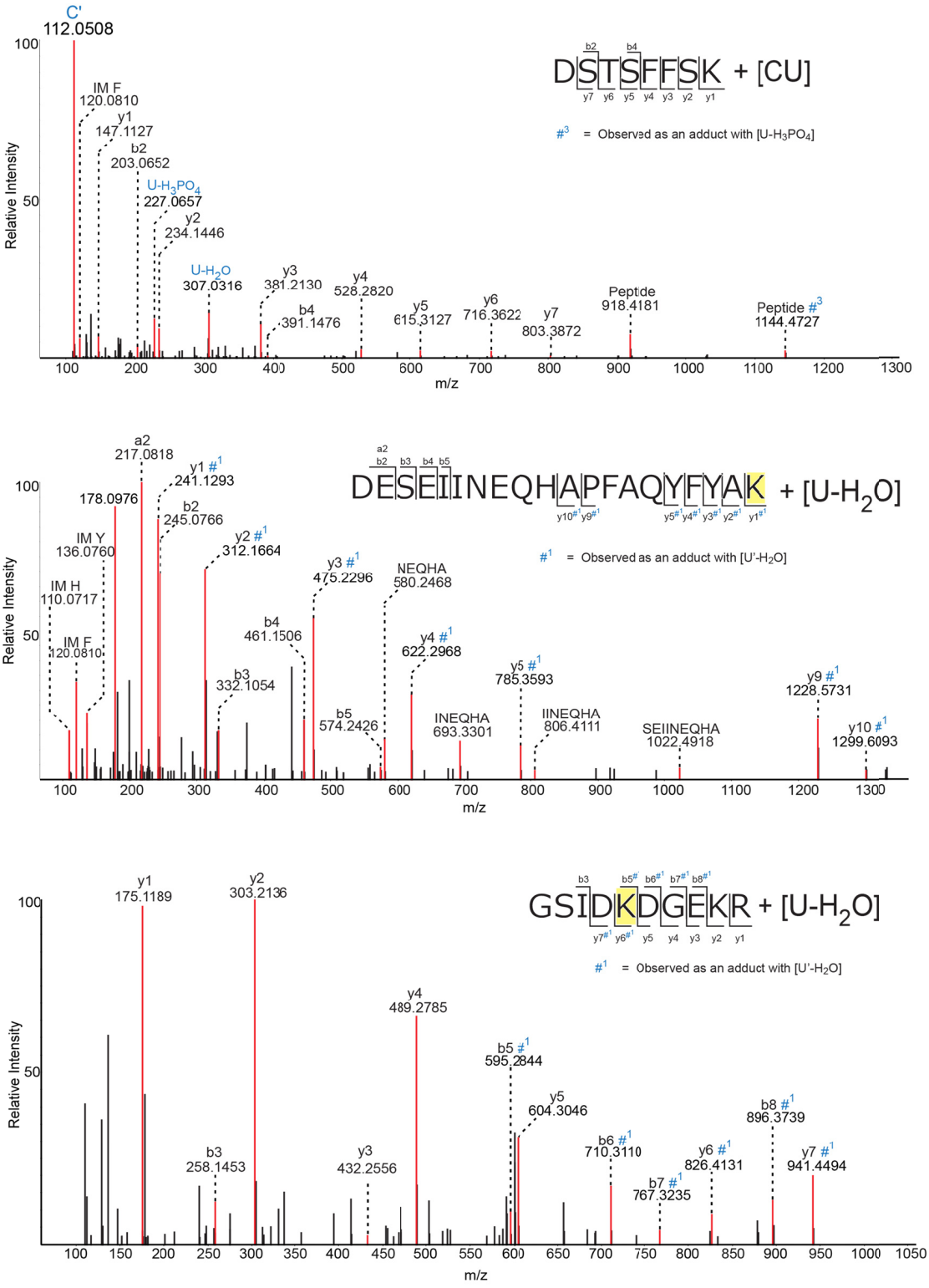




Supplementary Figure S6



Supplementary Figure S7



Supplementary Figure S8

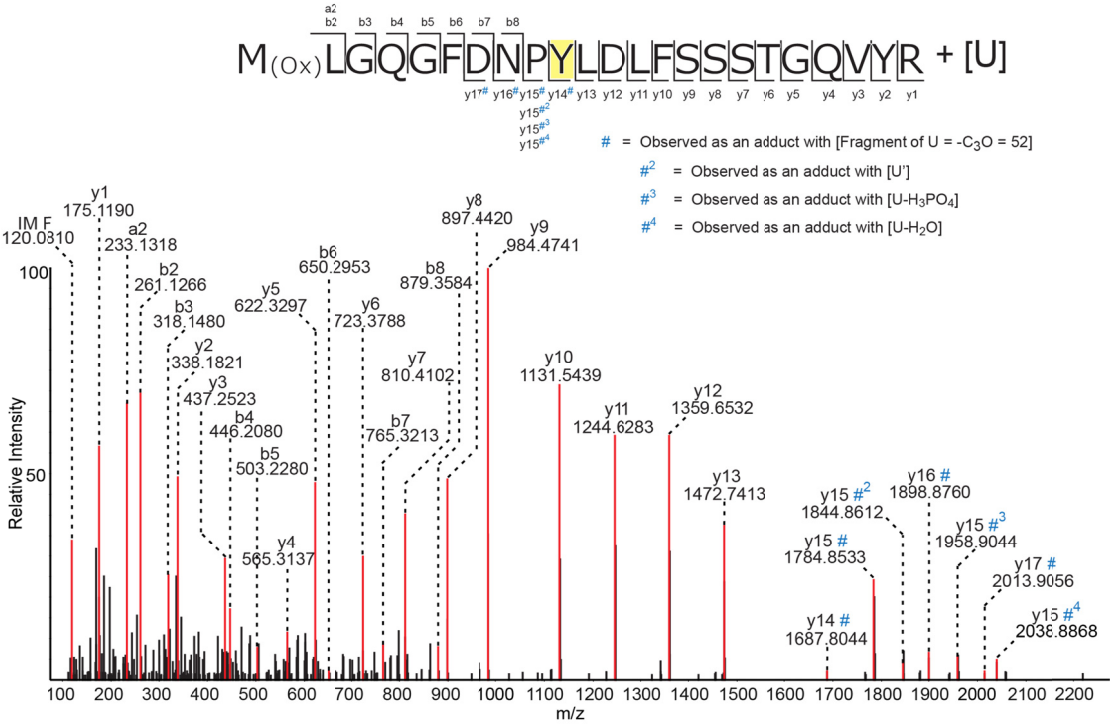


Fig. S4 – S8 MS/MS fragmentation spectra of the Protein – RNA crosslinks identified within the Type I-F Cascade complex from *Shewanella putrefaciens*

Shown are the sample spectra for every peptide observed to be cross-linked with an RNA moiety (see Fig. 3). In each spectrum, the cross-linked peptide sequence and its corresponding y- and b- type fragment ions are indicated at the top, referring to the ions which retain the charge on the N- or C- terminus, respectively. The cross-linked amino-acid residues are highlighted in yellow. Some of the b- and y- ions were observed with a mass shift of #, #1, #2, #3, #4 and #5 corresponding to -C3O (a fragment of Uracil), U'-H2O, U', U-H3PO4, U-H2O and U respectively. IM: Immonium ions; C': Base of C, 112.05 Da; U': Base of U, 112.0273 Da.

Supplementary Figure S9

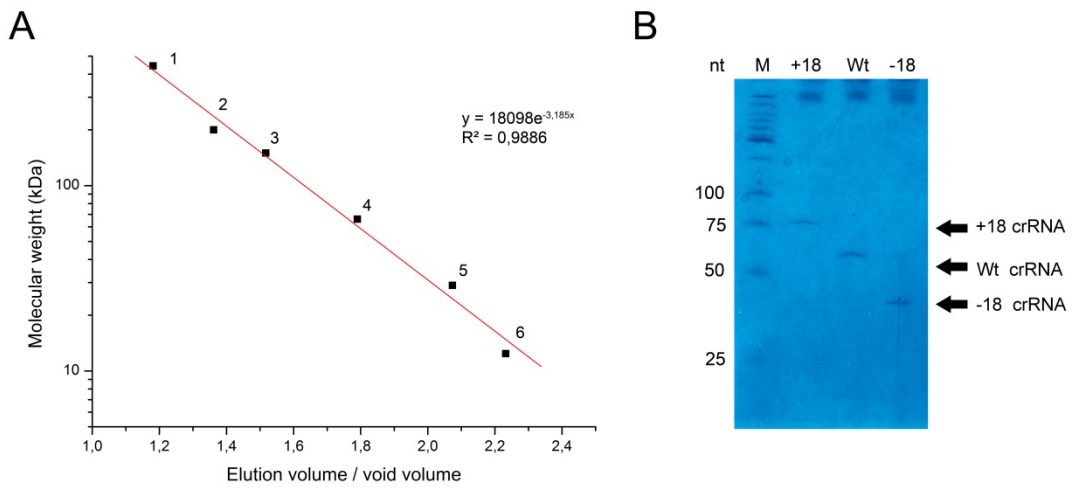


Fig. S9. Calibration of the size-exclusion column and analysis of crRNA maturation

(A) Calibration curve of the size-exclusion chromatography column Superdex 200 with the following protein standards: 1) apoferritin (443 kDa), 2) ß-amylase (200 kDa), 3) alcohol dehydrogenase (150 kDa), 4) bovine serum albumin (66 kDa), 5) carbonic anhydrase (29 kDa) and 6) cytochrome C (12.4 kDa). Elution volumes of these proteins were divided by the column's void volume (45 ml).

(B) 8 M urea PAGE and toluidine blue staining of purified Cascade variants with extended (+18), wildtype (wt) and shortened (-18) crRNA spacers. Processed crRNAs are identified and their length differences correspond to the modulated spacer length.

Supplementary Figure S10

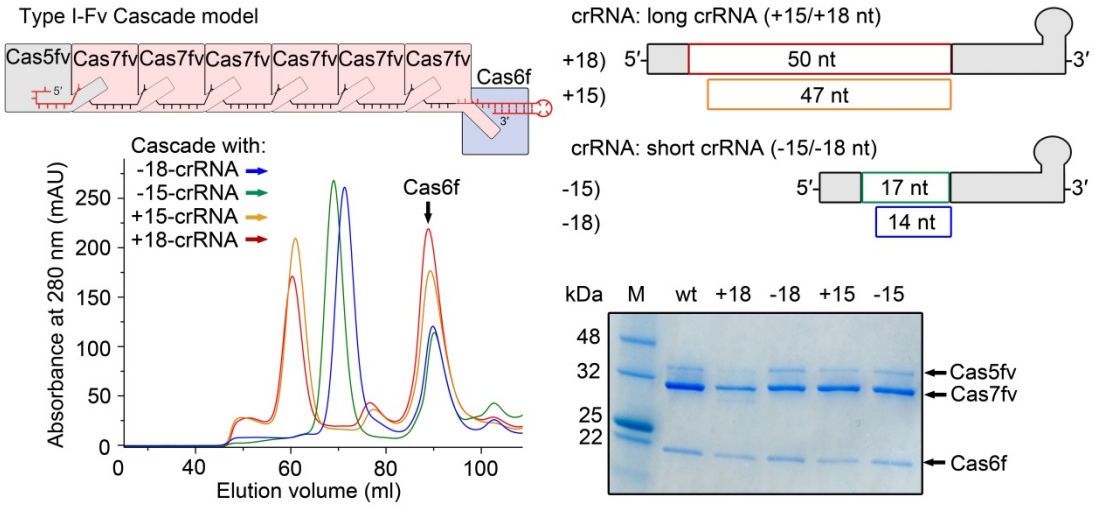


Fig. S10. Purification of recombinant synthetic Cascade variants with shortened and elongated spacers

Top left: Cartoon model of Type I-Fv Cascade. Cas5fv and Cas6f are proposed to cap the crRNA 5′- and 3′-repeat tags. Multiple subunits of Cas7fv are suggested to form a filamentous

backbone along the crRNA backbone in analogy to the Type I-E Cas7 backbone of E. coli. Additional large and small subunits are missing. Top right: Variants of the crRNAs wildtype spacer (32 nt) with short spacer length (14 & 17 nt) and long spacer length (50 & 47nt) were designed. Bottom left: Recombinant Cascade complexes were produced and purified via size-exclusion chromatography. Cas6f monomer peaks were observed. The relative shift of the major second peak during identical size-exclusion chromatography runs revealed efficient Cascade production and Cas7fv subunit number modulation in response to the crRNA's spacer length. Bottom right: SDS-PAGE analysis of purified Cascade variants confirms stable complex formation and reveals a varying Cas7fv band intensity in relation to Cas5fv and Cas6f band intensities.