Supplementary information for

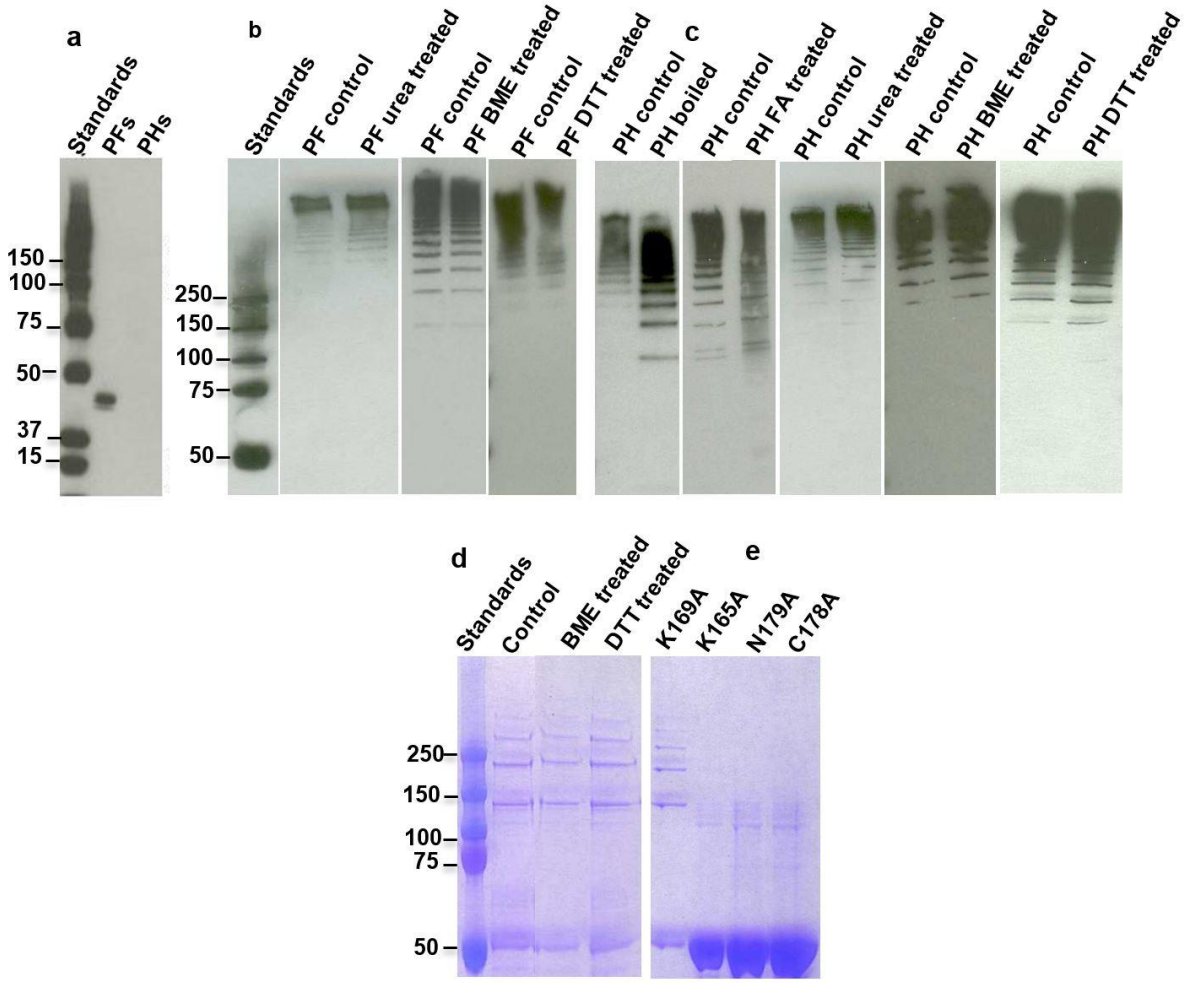
Spirochete flagellar hook protein self-catalyzes a lysinoalanine covalent cross-link for motility

Authors: Michael R. Miller, Kelly A. Miller, Jiang Bian, Milinda E. James, Sheng Zhang, Michael Lynch, Patrick S. Callery, Justin M. Hettick, Andrew Cockburn, Jun Liu, Chunhao Li, Brian R. Crane, and Nyles W. Charon*

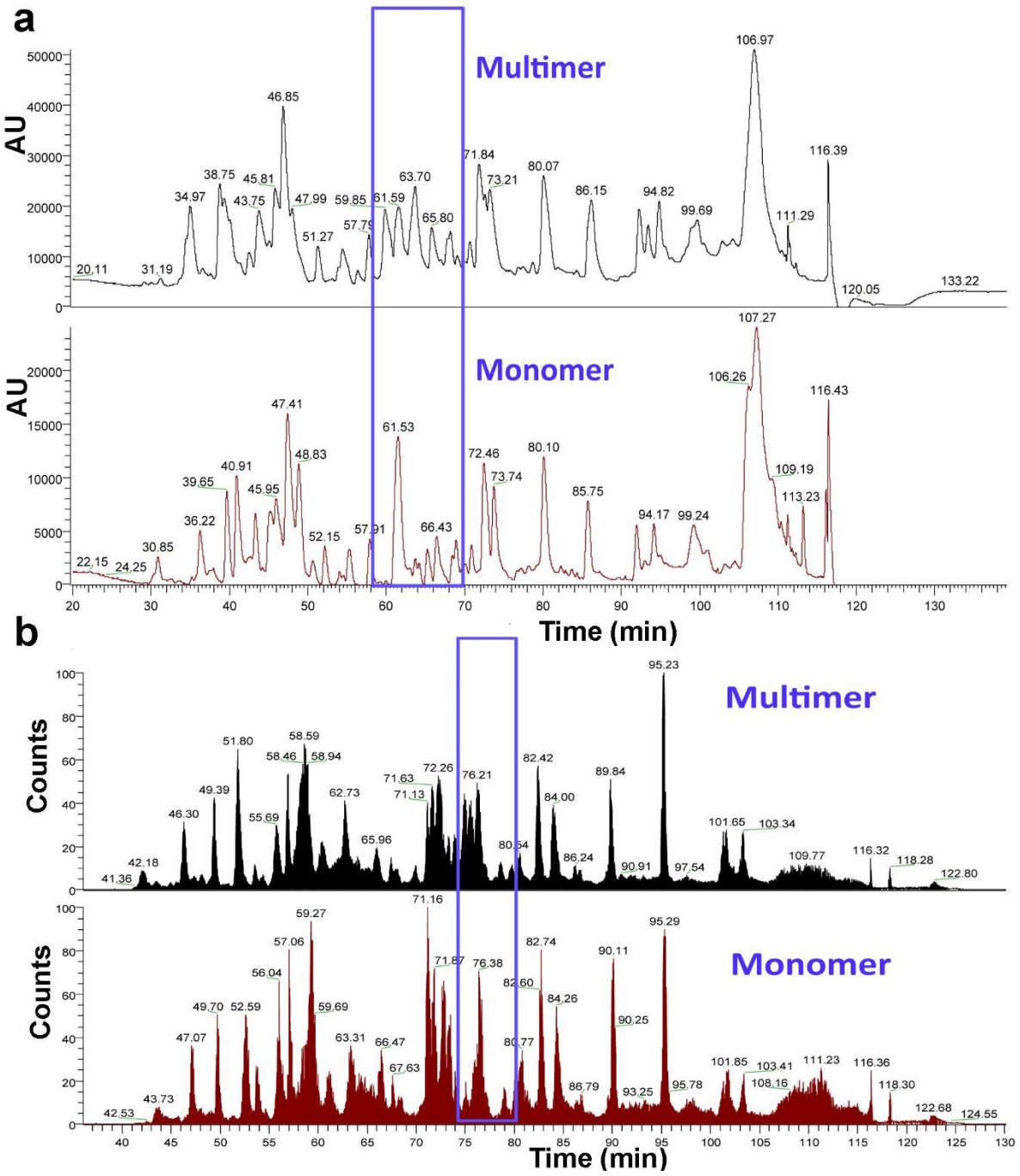
*Corresponding author. E-mail: ncharon@hsc.wvu.edu

This file includes:

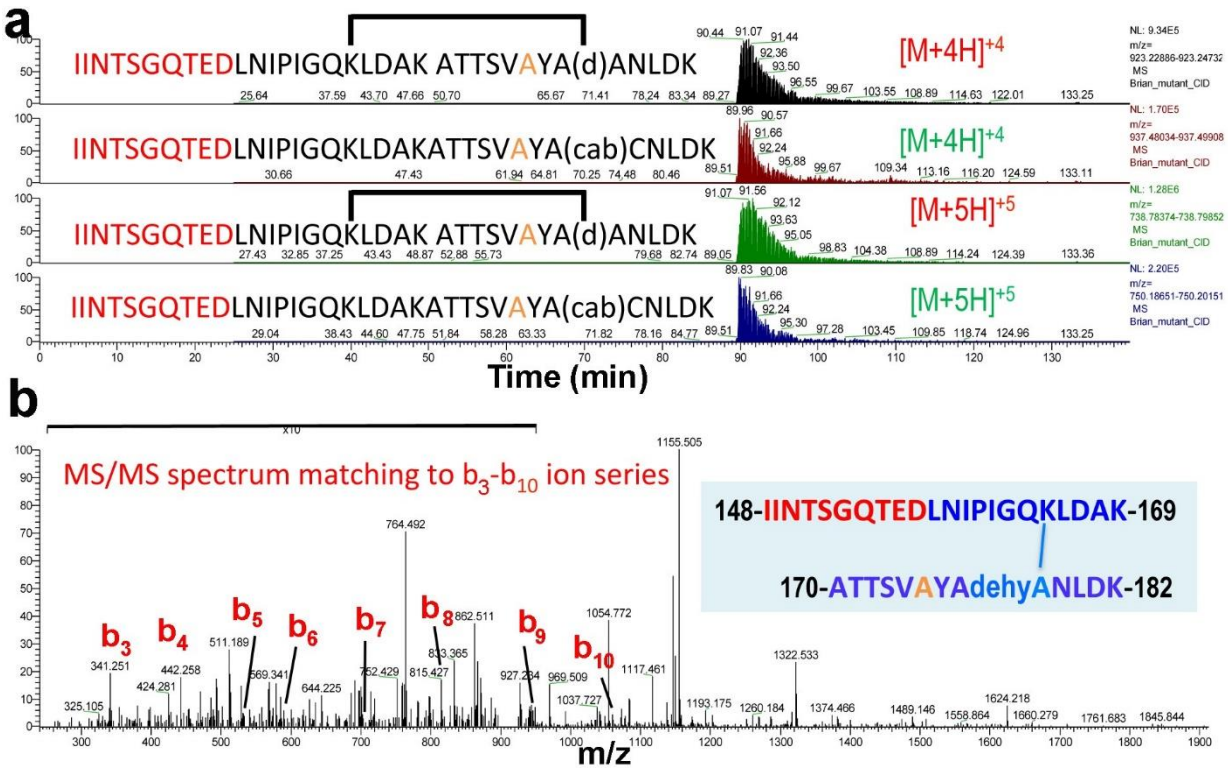
Supplementary Figures 1 – 12.
Supplementary Tables 1-6.
Captions for supplementary videos 1-4



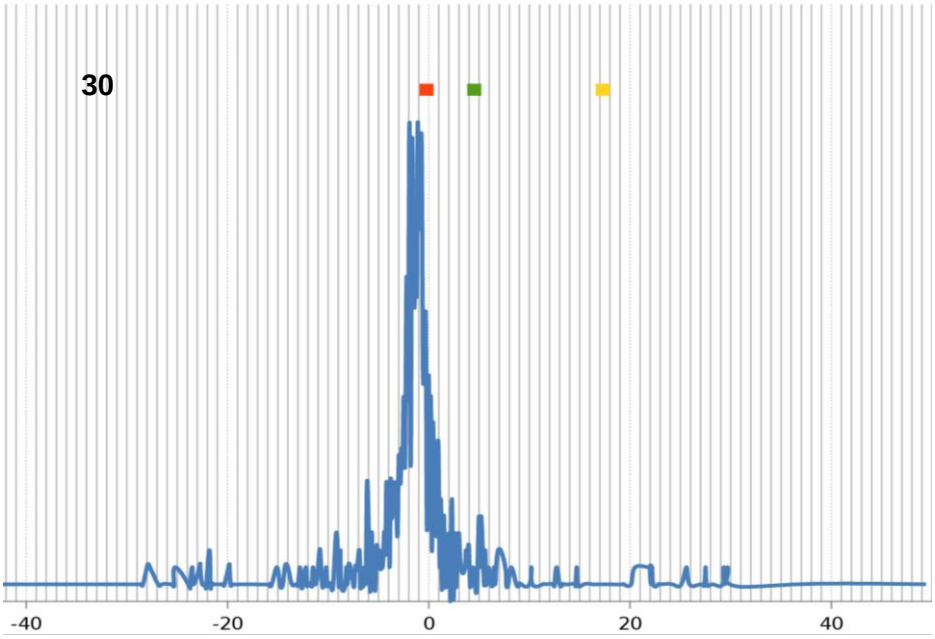
Supplementary Fig 1. Western blot and Imperial staining analysis of *T. denticola* periplasmic flagella (PFs), polyhooks (PHs), and *in vitro* synthesized high molecular weight complex (HMWC). (**a**) PFs and PHs were reacted with antibodies directed to FlaB. PFs (**b**) and PHs (**c**) were treated with various agents or boiled and analyzed by western blot with antibodies directed to FlgE: Formic acid (FA), β-mercaptoethanol (BME), dithiothreitol (DTT). Although some breakdown of the HMWC occurred with some treatments, no monomeric rFlgE was detected. (**d**) *In vitro* synthesized HMWC's were reacted with BME or DTT and analyzed by SDS-PAGE and Imperial stained. (**e**) *In vitro* synthesized HMWC's of four mutant proteins after one week incubation under cross-linking conditions. Note that K165A, N179A, and C178A fail to form HMWCs. Blots and gels were repeated at least 3 times, and the data shown are representative. Figure e was only done once.



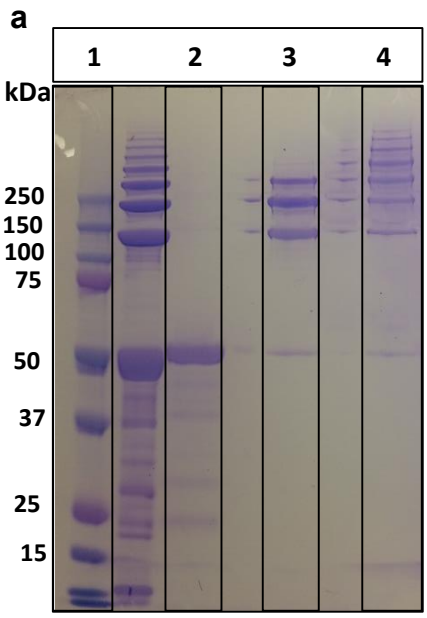
Supplementary Fig 2. LC-MS analysis of *T. denticola* rFlgE tryptic peptides (**a**) Comparison of UV elution profile taken from the monomer and multimer samples of *T. denticola* rFlgE. Differences in peptide absorbance are boxed. (**b**) MS total ion count profile of tryptic digests in (a). Regions of significant difference are boxed and were analyzed further for peptide identification. A major peptide of MW = 3732.9290 Da was identified only in the multimer sample. Note that there is 11 min offset between the UV and MS traces that was calibrated as described in the Methods Section.

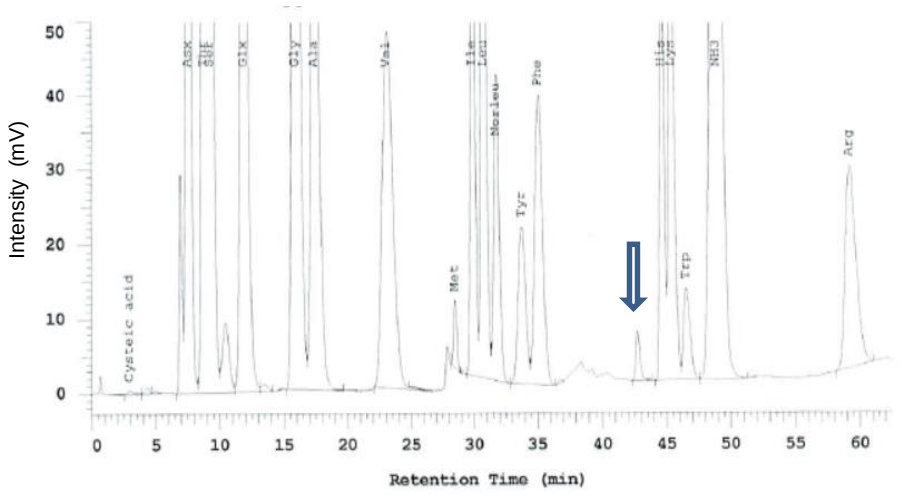


Supplementary Fig 3. MS/MS characterization of *T. denticola* rFlgE N175A substitution. (**a**) Representative extracted ion chromatograms (XICs) of the tryptic peptide derived from *in vitro* formed HMWC multimeric species reveal the expected molecular mass for the lysinoalanine crosslink between K165 and dehydroalanine 178 (d) in the derivative inter-peptide, thereby confirming the C-terminal sequence. The N175A mutation is identified in yellow. XIC's for both the cross-linked, and noncross-linked carbamidomethyl (cab)-modified peptides are shown. (**b**) Representative MS/MS spectrum matching to a b_3-b_{10} ion series.



Supplementary Fig. 4. Test of 3 models of cross-linking. MS data obtained from *in vitro* formed *T. denticola* rFlgE HMWC was examined to determine the accuracy of the Protea instrument and to test 3 models whereby T13-T14 is linked to T15, with the loss of 34 Da. The ppm errors of m/z values (M ion) for 1000 peptides automatically identified as FlgE by the software were plotted (blue line). The software was set to allow an error of up to 50 ppm, but it is clear from the graph that the instrument is accurate to within ~5 ppm. It also appears that it was biased to give a slight negative ppm error. In this run the observed mass unique to the HMWC was 3731.9209. This was compared to the predicted masses from the three models to determine the ppm error for each model [loss of one SH_2 (red), two NH_3's (yellow) or one H_2O_2 (green)], The loss of one SH_2 (3731.9218) provided an excellent fit; loss of one H_2O_2 (3731.8564) could not be eliminated, but did not fit as well; loss of two NH_3's was eliminated as a possibility, based on this analysis.

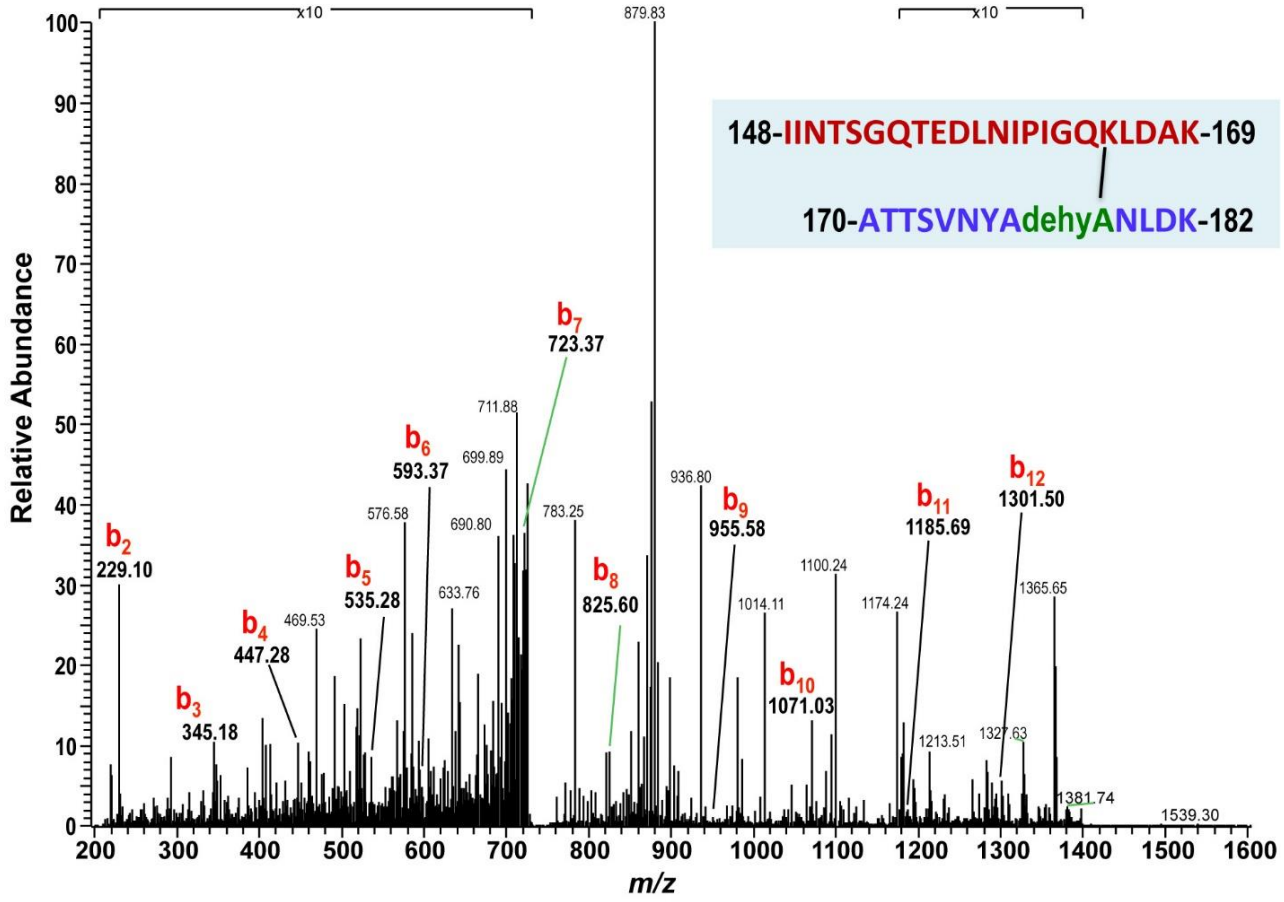




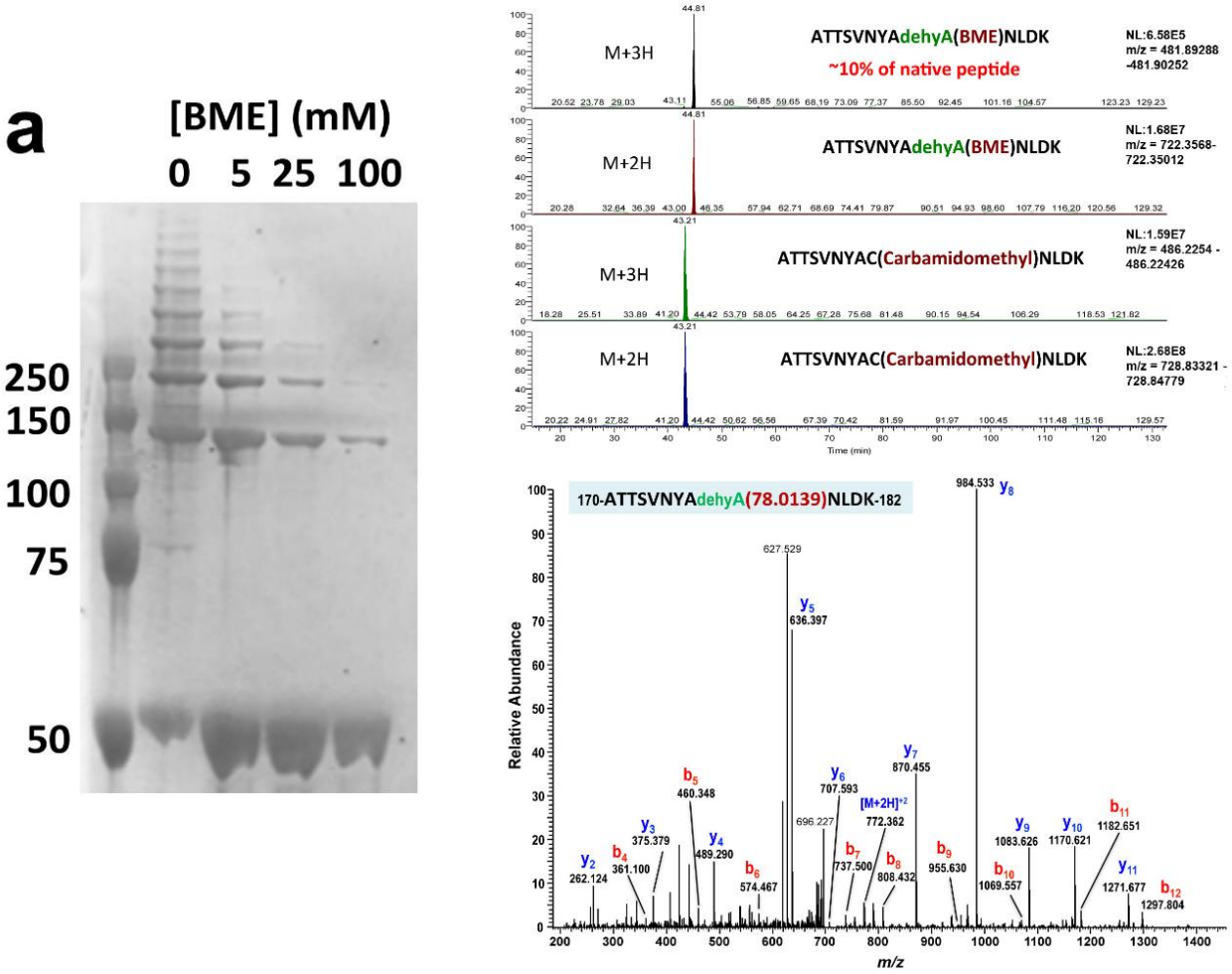
c

	Gly	Ala	Tyr
#AA/FlgE*	54	37	10
9LAL/10FlgE†	0.017	0.024	0.09
~50 kDa LAL/AA‡	0.006	0.011	0.043
~150 kDa LAL/AA	0.015	0.026	0.100
>250 kDa LAL/AA	0.021	0.035	0.140

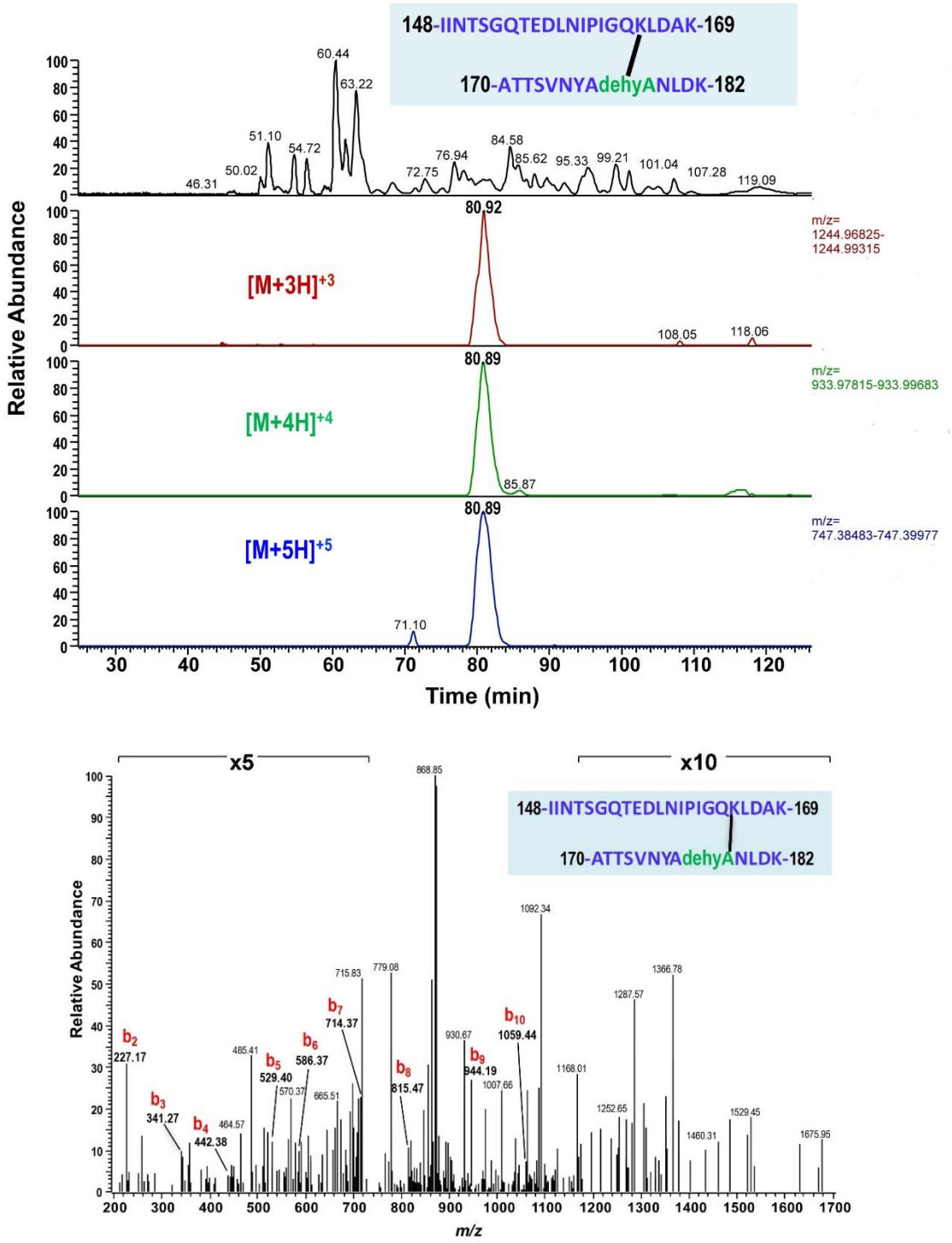
Supplementary Fig 5. Lysinoalanine (LAL) analysis of FlgE monomer and HMWC. *In vitro* crosslinked *T. denticola* rFlgE was electrophoresed in agarose, regions corresponding to monomer (50 kDa), ~150 kDa, and >250 kDa HMWCs were excised, and the protein electroeluted. (**a**) Electroeluted proteins were electrophoresed and Imperial stained (outlined lanes): 1 = dual color protein markers ; 2 = electroeluted 50 kDa monomer; 3 = electroeluted ~150 kDa HMWC; 4 = electroeluted ≥250 kDa HMWC. Proteins were submitted for amino acid analysis: (**b**) Partial amino acid chromatogram of ≥250 kDa HMWC; arrow indicates the position of LAL, well resolved from other amino acids. (c) *The # of Gly, Ala,Tyr in one FlgE . †The expected ratio of LAL to Gly, Ala and Tyr in a HMWC of 10 FlgEs joined by 9 LALs. ‡Ratios of LAL to Gly, Ala and Tyr determined in the 3 electroeluted protein fractions from one Td rFlgE preparation. The levels of LAL in the HMWCs were similar to that in a complex of 10 FlgE monomers joined by 9 LALs. LAL levels were much lower in the monomer, likely arising from dehydroalannine in the monomer (Supplementary Fig. 7) forming intra-molecular LAL cross-links. In four Td rFlgE preparations, the level of LAL was 3-8 times higher (mean of 5) in the >250 kDa HMWC than in the monomer.



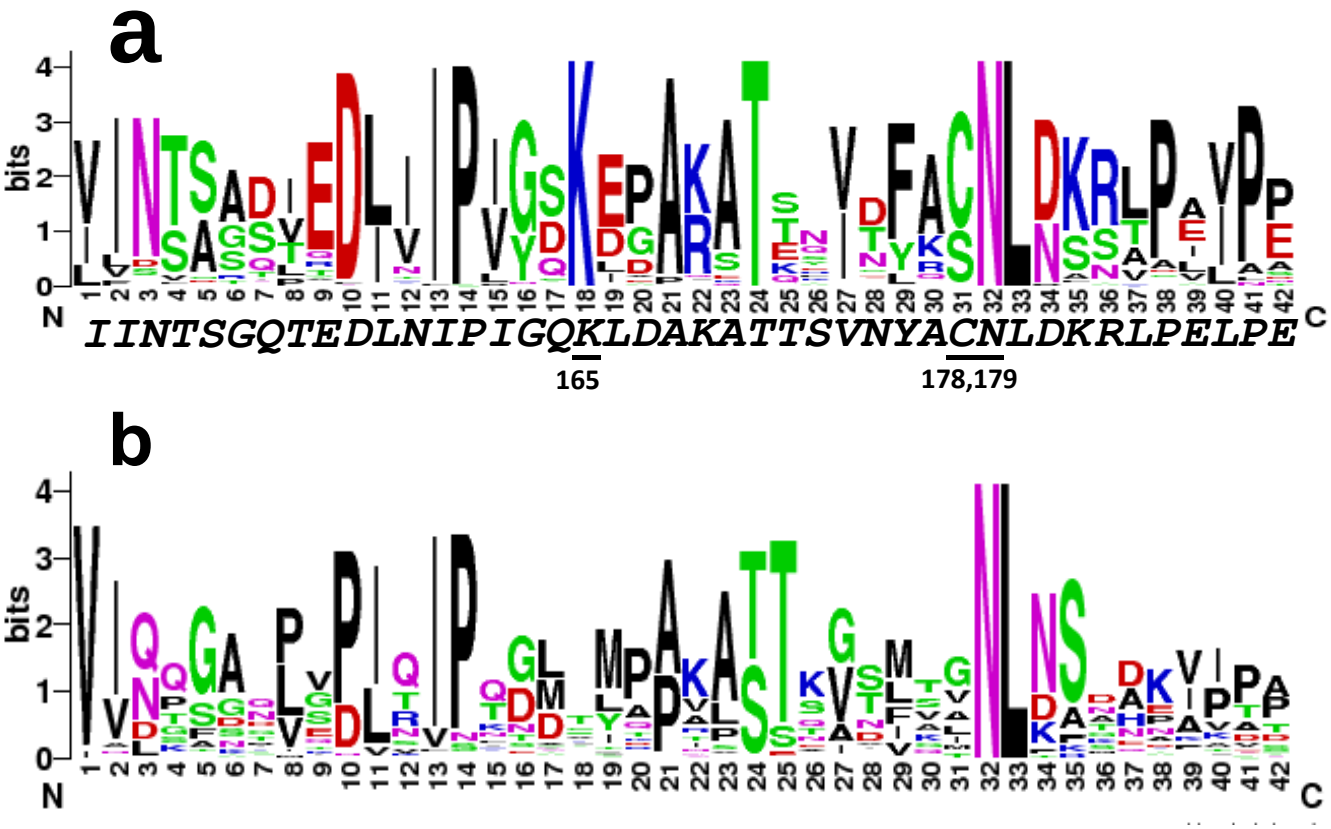
Supplementary Fig. 6. MS/MS spectrum of isotopically mixed *T. denticola* cross-linked peptide. Representative MS/MS spectrum of the 752.9752^{+5} peptide representing a ^{15}N (red) - ^{14}N (blue) cross-link as indicated in the inset. Mass corresponds to that expected for the LAL cross-link, with a b-ion series extending to N-terminal residue 12. Mass peaks at low and high m/z ratios are shown at 10x amplitude.



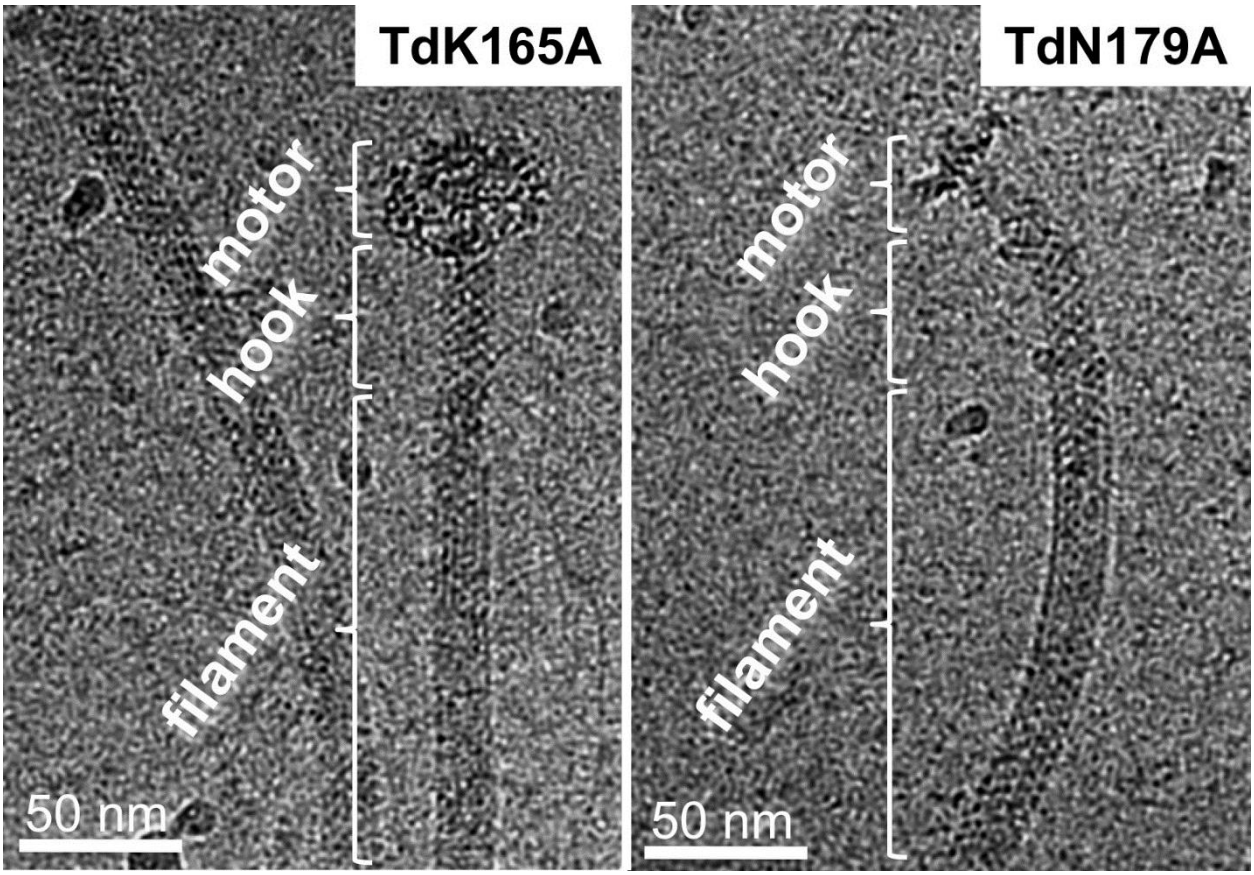
Supplementary Fig. 7. BME blocks cross-linking by adding to dehydroalaine. (**a**) Inhibition of *T. denticola* rFlgE cross-link formation by addition of BME as shown in a representative gel. (**b**) MS analysis: Above: XIC of tryptic digests from momeric rFlgE treated with BME to inhibit cross-linking. ~10% of the detected peptides contain a BME adduct with dehydroalanine (Cα-CH_2-S-$(CH_2)_2$-OH). The remaining peptides retain Cys at position 178 and thus undergo modification by iodoacetamide to form a carbamidomethyl modification. Below: MS/MS spectrum of m/z = 722.3429^{2+} identifying the target peptide with an adduct between BME and dehydroalanine in the form of a thio-ether linkage



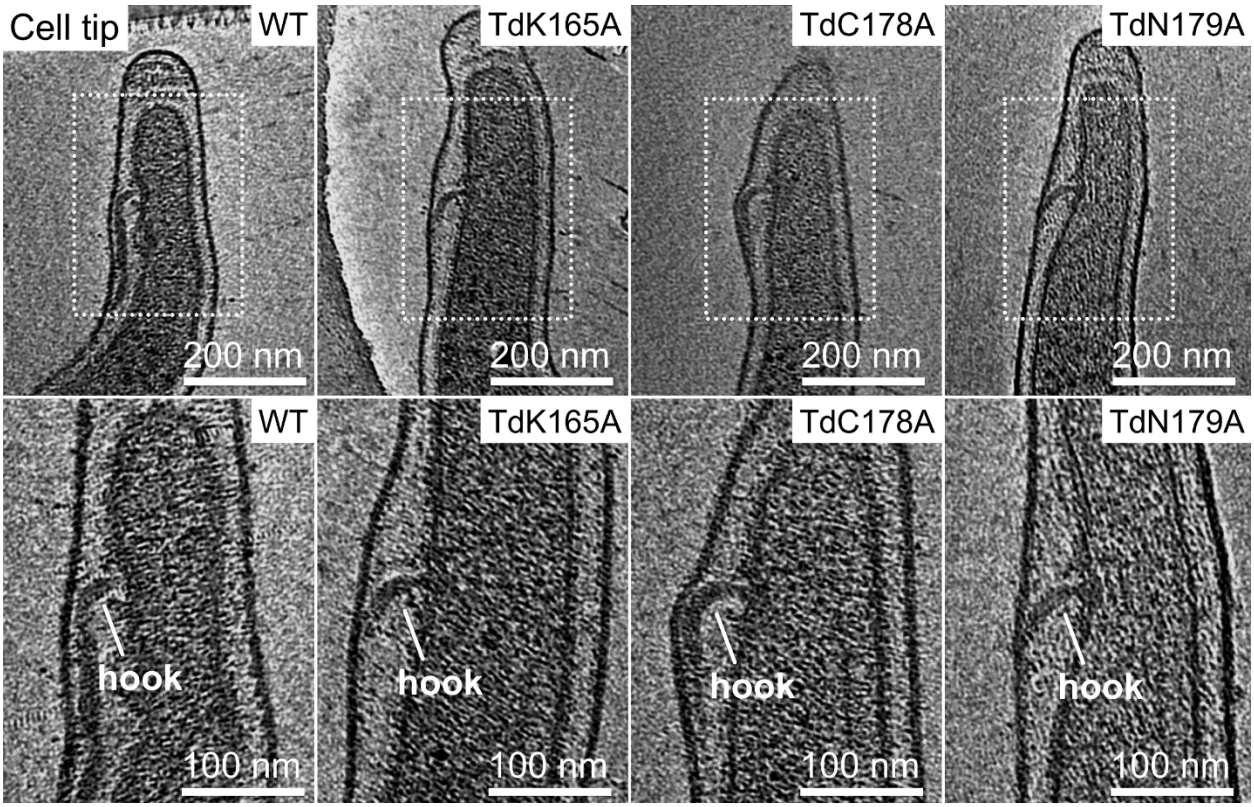
Supplementary Fig. 8. The cross-linked peptide is highly represented in flagellar PHs extracted from Td. **a)** Representative Extracted Ion Chromatograms (XICs) of the cross-linked peptide in Td PHs. XICs are shown for the $[M+nH]^{n+}$ n=3,4,5 tryptic peptides from the PHs. Top trace shows the total ion counts for all chromatographed peptides vs. LC retention time (min). **b**) MS/MS spectrum of 747.39142^{+5} cross-linked peptide of Td PH FlgE. Mass peaks at low and high m/z ratios are shown at 5x or 10x amplitude.



Supplementary Fig. 9. Conserved residues in the spirochete cross link region. (**a**) Phylogenetic tree analysis was carried out across 176 unique bacterial FlgE sequences based on alignment of Td peptide: IINTSGQTEDLNIPIGQKLDAKATTSVNYACNLDKRLPELPE. Conservation of the cross-linking peptide distinguishes two families whose sequence conservation via WebLogo is shown. In spirochetes and the closely related Synergistetes (92 sequences), Lys165 is invariant, and position 178 is either Cys, or sometimes Ser, two residues that are both known to participate in lysinoalanine formation. Asn179 and Leu180 are also invariant in this family. The Td amino acid sequence is shown in italics, with those residues essential for cross-linking underlined and numbered below. (**b**) Other types of bacteria (84 sequences) do not conserve the cross-linking residues in their FlgE sequences, but do conserve Asn179 and Leu180.



Supplementary Fig.10. Cryo-EM of TdK165A and TdN179A. PFs were purified from the above two mutants and TdC178A (Fig. 4c) that fail to cross-link FlgE and have altered motility, and examined by cryo-EM. These structures were compared to that of the WT (Fig. 1a), and representative Cryo-EM images are shown. No discernable differences in the flagellar hook and filament structure were detected among the mutants compared to the WT.



Supplementary Figure 11. Cryo-EM reconstructions of wild-type and mutants. Top panels show tomographic slides from the cell tips of WT, TdK165A, TdC178A, TdN179A mutants, respectively. The bottom panels show the zoom-in views from the area highlighted in the corresponding top panel. Both flagella and hooks remain intact in periplasmic space in all strains. Representative Cryo-EM reconstructions are presented.

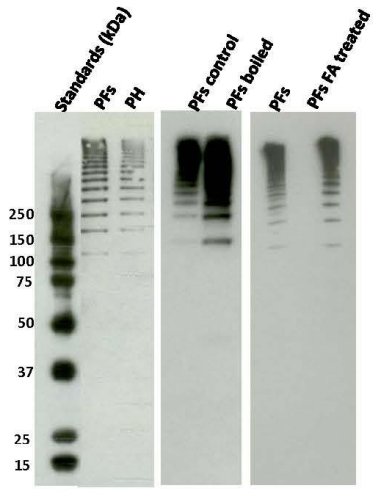


Fig 1B

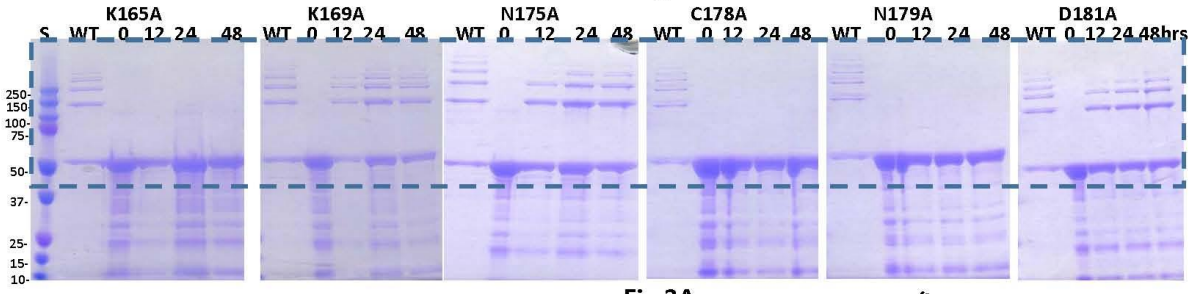


Fig 3A

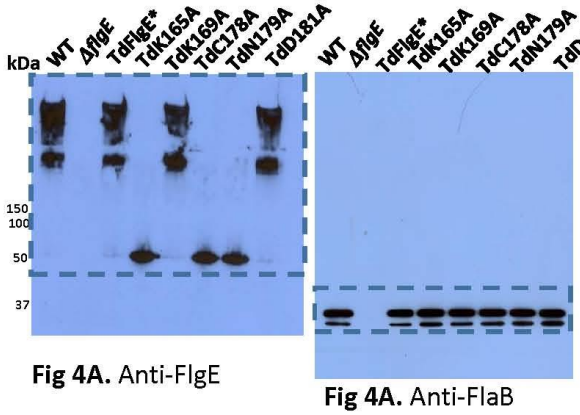


Fig 4A. Anti-FlgE

Fig 4A. Anti-FlaB

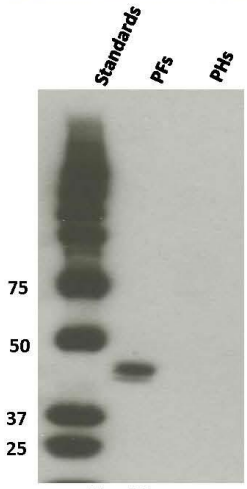


Fig S1a

Supplementary Figure 12. Uncropped gels and blots.

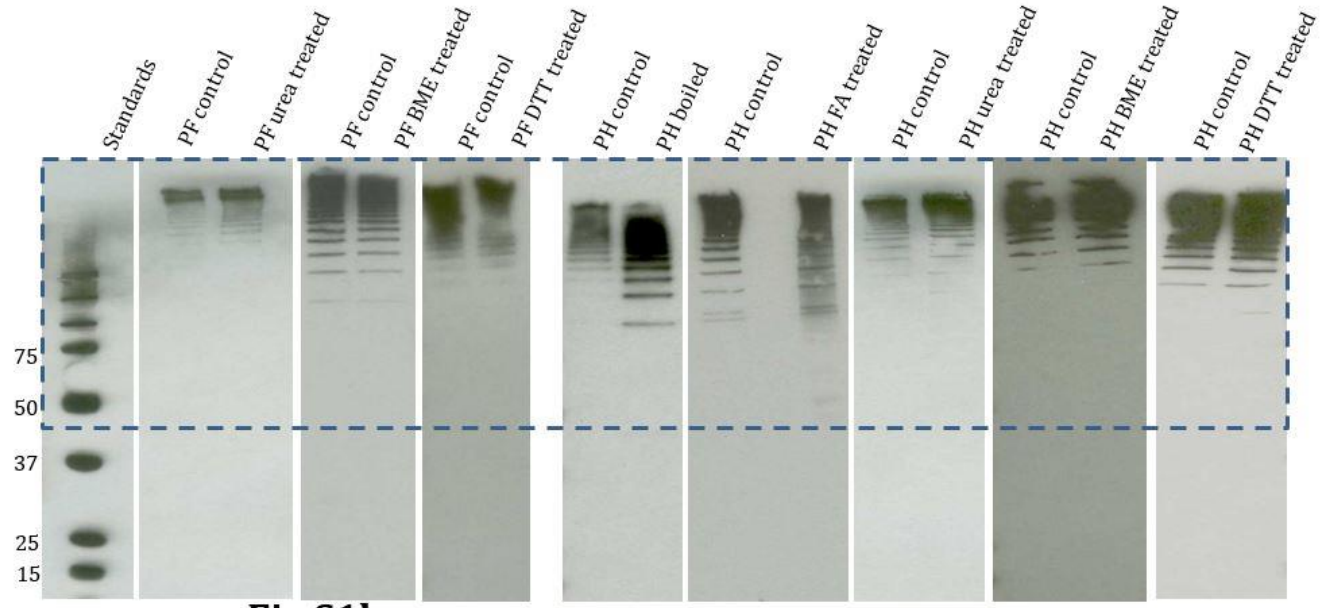


Fig S1b **Fig S1c**

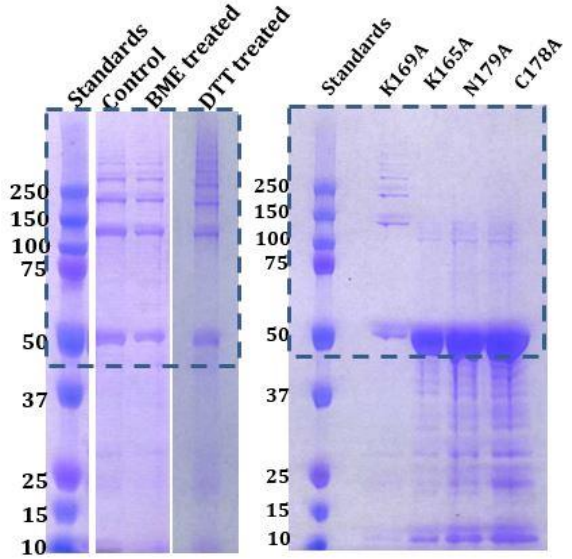


Fig S1d,e

Supplementary Fig 12 (continued). Uncropped gels and blots.

Supplementary Table 1. T13, T13-T14 and T15 Peptide abundances. Tryptic peptides involved in lysinoalainine formation were quantified by MS extracted ion chromatograms (XICs) in their uncrosslinked forms for monomeric FlgE and the HMWC. The uncrosslinked peptides in their various forms were standardized against two peptides that do not participate in cross-linking (red and blue sequences). See Supplementary Table 2 for summary.

Peptide	Residues	Peptide Sequence	Modification	Monomer Peak Area	HMWC Peak Area
1	170-182	ATTSVNYAcNLDK	C9 - Carbamidomethyl	7.92E+09	8.46E+09
2	170-182	ATTSVNYAcnLDK	C9- Carbamidomethyl; N10 - Deamidated	1.13E+08	7.51E+07
3	170-183	ATTSVNYAcNLDKR	C9 - Carbamidomethyl	5.12E+08	5.28E+08
	Σ			**8.55E+09**	**9.06E+09**
4	148-165	IINTSGQTEDLNIPIGQK		4.02E+09	2.93E+09
				1.31E+10	1.63E+10
5	148-165	IInTSGQTEDLNIPIGQK	N3 - Deamidated	3.94E+09	2.89E+09
				1.38E+10	1.70E+10
6	148-169	IINTSGQTEDLNIPIGQKLDAK		3.00E+07	2.50E+07
7	148-182	IINTSGQTEDLNIPIGQKLDAKATTSVNYAcNLDK	C31 - Carbamidomethyl	3.08E+07	3.40E+07
	Σ			**3.49E+10**	**3.91E+10**
	194-206 (std1)	**AQILESTWSTEFK**		7.90E+09	2.28E+10
				1.02E+08	8.30E+08
	Σ			**8.00E+09**	**2.36E+10**
	706-722 (std2)	**VYDSFGEAHELQIDFAR**		9.32E+09	2.53E+10
				1.70E+09	6.76E+09
	Σ			**1.10E+10**	**3.21E+10**

Supplementary Table 2. Non-crosslinked peptide ratios in monomer compared to HMWC. Summary of peptide abundance data presented in Supplementary Table 1.

[a] – standardized to std1 (red), [b] – standardized to std2 (blue)

Peptides	Monomer[a]	HMW[a]	Monomer/ HMW[a]	Monomer[b]	HMW[b]	Monomer/ HMW[b]	Average of Std1 and Std2
1-3	1.0679	0.3835	**2.78**	0.7758	0.2823	**2.75**	**2.77**
4-7	4.3581	1.6564	**2.63**	3.1658	1.2191	**2.6**	**2.62**

Supplementary Table 3. Recombinant *Treponema denticola* FlgE Unique Peptides. Native peptides identified by MS/MS from trypsin digest of recombinant Td FlgE. Sequence coverage = 89.4%.

Measured MR	Measured MR	Z	Calculated MR	ΔMR	Missed Sites	Mascot Score	Score Threshold	Expect Value	Sequence
339.6819	677.3493	2	677.3497	-0.0003	0	40.89	49	7.00E-02	GFQAGAK
344.1772	686.3398	2	686.3388	0.0011	0	23.27	40	5.10E-01	TFYTR
389.7087	777.4028	2	777.4021	0.0007	0	47.86	48	1.20E-02	AGAFGIDK
514.2701	1026.5256	2	1026.5247	0.0009	0	61.05	46	3.20E-04	VPGEVNAWR
527.2645	1052.5144	2	1052.5138	0.0005	0	57.72	47	9.50E-04	NTITQFSDK
548.2929	1094.5711	2	1094.572	-0.0008	0	55.24	50	3.20E-03	LPELPEGANR
626.3448	1250.6751	2	1250.6731	0.0021	1	42.56	49	5.50E-02	RLPELPEGANR
629.8146	1257.6146	2	1257.6136	0.0011	0	53.97	48	3.00E-03	EGTLVNPANGMR
728.8405	1455.6663	2	1455.6664	0	0	71.71	47	3.90E-05	ATTSVNYACNLDK
755.3419	1508.6692	2	1508.6718	-0.0026	0	87.7	45	6.50E-07	VQGWMAEEAEGFR
759.3827	1516.7508	2	1516.7522	-0.0014	0	102.98	50	6.30E-08	HTFDVNLGEIGTSK
509.2302	1524.6687	3	1524.6667	0.002	0	68.22	45	5.20E-05	VQGWMAEEAEGFR[a]
770.3886	1538.7627	2	1538.7617	0.001	0	97.56	51	2.50E-07	AQILESTWSTEFK
538.263	1611.7672	3	1611.7675	-0.0003	1	46.41	50	2.60E-02	ATTSVNYACNLDKR
808.8972	1615.7798	2	1615.7802	-0.0004	0	101.2	50	8.70E-08	ATVNVDPTNADATATR
831.9042	1661.7939	2	1661.7944	-0.0005	0	122.84	50	5.90E-10	SLFSGVTGMQNHQTR
831.9432	1661.8719	2	1661.8737	-0.0018	0	128.11	53	3.50E-10	VGIGTTDGVQNSFIVR
832.3975	1662.7804	2	1662.7784	0.002	0	101.58	50	7.00E-08	SLFSGVTGMQNHQTR[b]
832.4362	1662.8579	2	1662.8577	0.0002	0	120.89	53	1.80E-09	VGIGTTDGVQNSFIVR
839.9016	1677.7887	2	1677.7893	-0.0007	0	112.87	49	5.10E-09	SLFSGVTGMQNHQTR[a]
839.942	1677.8695	2	1677.8686	0.0009	0	82.61	53	1.20E-05	IDQSGIITGVYSNGVR
897.4451	1792.8757	2	1792.8778	-0.0021	0	104.35	52	6.40E-08	MDVIGNNVANVNTTGFK
903.98	1805.9455	2	1805.9445	0.0011	0	133.14	54	1.20E-10	TIQTSDTMLETVLNLK
969.9473	1937.8801	2	1937.8795	0.0006	0	117.75	48	1.30E-09	AYEQDGYTLGYLENFR
971.0204	1940.0262	2	1940.0215	0.0048	0	108.98	53	3.00E-08	IINTSGQTEDLNIPIGQK
654.6644	1960.9713	3	1960.9676	0.0036	0	118.15	53	3.50E-09	QEIGQIAMAGFANQGGLEK
655.0237	1962.0492	3	1962.0456	0.0036	1	37.57	54	4.80E-01	TIQTSDTMLETVLNLKR
998.9734	1995.9323	2	1995.9326	-0.0003	0	110.69	50	1.10E-08	VYDSFGEAHELQIDFAR
1227.5927	2453.1707	2	2453.1783	-0.0075	0	152.15	53	1.40E-12	AGQNTYVQSNNSGIANVSTSGTVGK
956.8352	2867.4838	3	2867.4777	0.0061	0	101.36	57	3.90E-07	VNFQDLISQQLSGAARPTEELGGVNPK
1519.231	3036.4474	2	3036.4461	0.0013	0	126.78	55	7.10E-10	GYFIGGTLEMSNVDLTDQFVDMIVTQK[b]
1207.945	3620.813	3	3620.8257	-0.0127	1	93.22	58	3.60E-06	IDQSGIITGVYSNGVRQEIGQIAMAGFANQGGLEK
1508.7397	4523.1974	3	4523.1953	0.0021	0	105.91	59	2.10E-07	FDNNGHLASVTDTAGNVTSPAGQVLVQISYNVVGANPDEAGAPTR
344.1772	686.3398	2	686.3388	0.0011	0	23.27	40	5.10E-01	TFYTR
389.7086	777.4027	2	777.4021	0.0006	0	43.47	48	3.20E-02	AGAFGIDK
548.2931	1094.5717	2	1094.572	-0.0002	0	49.19	48	8.80E-03	LPELPEGANR
582.8038	1163.5931	2	1163.5935	-0.0004	0	65.45	51	4.10E-04	NVANVNTTGFK
626.3445	1250.6744	2	1250.6731	0.0013	1	53.97	48	5.00E-02	RLPELPEGANR
629.8146	1257.6146	2	1257.6136	0.0011	0	39.2	48	9.10E-02	EGTLVNPANGMR
728.8405	1455.6665	2	1455.6664	0.0001	0	71.79	47	3.80E-05	ATTSVNYACNLDK
755.3415	1508.6684	2	1508.6718	-0.0034	0	85.64	45	1.00E-06	VQGWMAEEAEGFR
770.3879	1538.7613	2	1538.7617	-0.0003	0	97.53	51	2.70E-07	AQILESTWSTEFK

971.0179	1940.0212	2	1940.0215	-0.0002	0	96.42	53	5.40E-07	IINTSGQTEDLNIPIGQK
956.8352	2867.4838	3	2867.4777	0.0061	0	101.36	57	3.90E-07	VNFQDLISQQLSGAARPTEELGG VNPK
339.6821	677.3496	2	677.3497	0	0	40.69	49	7.40E-02	GFQAGAK
728.8408	1455.667	2	1455.6664	0.0006	0	68.14	47	8.70E-05	ATTSVNYACNLDK
538.263	1611.7672	3	1611.7675	-0.0003	1	46.41	50	2.60E-02	ATTSVNYACNLDKR
903.98	1805.9455	2	1805.9445	0.0011	0	133.14	54	1.20E-10	TIQTSDTMLETVLNLK
655.0236	1962.049	3	1962.0456	0.0035	1	38.11	54	4.10E-01	TIQTSDTMLETVLNLKR
495.7407	989.4669	2	989.4665	0.0004	0	61.59	46	3.30E-04	IEDTDIER
510.6003	1528.7792	3	1528.7787	0.0005	0	38.81	51	1.80E-01	FAPSPTGYLHVGGAR
407.2632	812.5119	2	812.512	-0.0001	0	72.63	43	1.20E-05	LAAQAIVK
405.2238	808.4331	2	808.433	0	0	51.86	47	3.70E-03	LASYIDK
400.1901	798.3656	2	798.366	-0.0005	0	53	44	1.40E-03	QGFYER
760.3917	1518.7688	2	1518.7599	0.0088	0	49.37	52	2.00E-02	AGVSMDEIIEQLSK
747.0507	2238.1303	3	2238.146	-0.0156	2	32.97	54	1.60E+00	RRSHYHWLHFQFLNAAR
546.7694	1091.5243	2	1091.5281	-0.0038	1	48.27	48	1.00E-02	AMNEKGVAEK[a]
420.7505	839.4864	2	839.4865	-0.0001	0	44.88	42	5.40E-03	QPLQVVR[b]
499.3113	996.608	2	996.608	0	0	42.08	40	7.30E-03	VALAAVVQAR
513.2399	1024.4652	2	1024.4648	0.0004	0	31.22	46	3.00E-01	YAVGAECAGK
499.7015	997.3885	2	997.3876	0.0009	0	35.63	36	1.30E-02	LNYDDDDK[c]
383.7086	765.4027	2	765.4021	0.0006	0	36.74	46	1.00E-01	VNTASFK
541.2493	1080.4841	2	1080.4757	0.0084	0	28.78	45	4.50E-01	AVSLTMDQSN[a]
771.3379	1540.6613	2	1540.665	-0.0036	0	25.76	44	6.80E-01	IGGMDAYQAMADAAR[b]

[a] - Methionine oxidation, [b] - NQ deamidation, [c] - 2-oxidation **MZ** - Measured m/z ratio; **MR** – Average charge-adjusted mass; **Missed Sites** – uncleaved trypsin sites in peptide; **Mascot Score** – Statistical measure of sequence match based on MS/MS parent and fragment ion patterns; **Score Threshold** – Threshold above which Mascort Score will give a match $P > 0.05$ considering all peptides in Mascot data base. **Expect Value** – expectation value – number of matches with equal or better scores that would be expected to occur by chance.

Supplementary Table 4. ***Treponema denticola*** **PH Unique Peptides.** Native peptides identified by MS/MS from trypsin digest of Td poly-hooks. Sequence coverage = 83.2%

Measured MZ	Measured MR	Z	Calculated MR	ΔMR	Missed Sites	Mascot Score	Score Threshold	Expect Value	Sequence
514.2701	1026.5256	2	1026.5247	0.0009	0	61.84	49	2.40E-03	VPGEVNAWR
527.2643	1052.514	2	1052.5138	0.0002	0	58.73	49	6.20E-03	NTITQFSDK
548.2933	1094.572	2	1094.572	0	0	46.35	50	1.50E-02	LPELPEGANR
626.3438	1250.6731	2	1250.6731	0	1	43.66	50	2.10E-02	RLPELPEGANR
629.8146	1257.6146	2	1257.6136	0.0011	0	54.65	50	9.80E-03	EGTLVNPANGMR
728.8406	1455.6667	2	1455.6664	0.0003	0	80.91	49	2.70E-06	ATTSVNYACNLDK
755.3441	1508.6736	2	1508.6718	0.0018	0	90.68	49	2.50E-07	VQGWMAEEAEGFR
759.3832	1516.7519	2	1516.7522	-0.0003	0	105.89	53	3.20E-09	HTFDVNLGEIGTSK
770.3884	1538.7622	2	1538.7617	0.0005	0	95.26	52	1.20E-07	AQILESTWSTEFK
808.8976	1615.7807	2	1615.7802	0.0005	0	121.13	52	6.30E-09	ATVNVDPTNADATATR
831.9044	1661.7943	2	1661.7944	-0.0001	0	106.02	52	2.30E-08	SLFSGVTGMQNHQTR
831.944	1661.8735	2	1661.8737	-0.0002	0	131.93	53	1.50E-10	VGIGTTDGVQNSFIVR
555.2653	1662.774	3	1662.7784	-0.0045	0	48.18	51	7.30E-03	SLFSGVTGMQNHQTR
905.4442	1808.8739	2	1808.8727	0.0012	0	101.28	53	7.70E-07	MDVIGNNVANVNTTGFK[a]
897.4462	1792.8779	2	1792.8778	0.0001	0	95.41	53	3.10E-06	MDVIGNNVANVNTTGFK[b]
903.98	1805.9454	2	1805.9445	0.0009	0	136.41	52	1.20E-10	TIQTSDTMLETVLNLK
969.9478	1937.8811	2	1937.8795	0.0015	0	105.25	50	7.70E-09	AYEQDGYTLGYLENFR
971.019	1940.0235	2	1940.0215	0.0021	0	108.37	52	5.50E-08	IINTSGQTEDLNIPIGQK
981.4923	1960.9701	2	1960.9676	0.0024	0	118.89	53	6.80E-10	QEIGQIAMAGFANQGGLEK
998.9741	1995.9336	2	1995.9326	0.0009	0	91.84	52	9.30E-07	VYDSFGEAHELQIDFAR
1227.5991	2453.1837	2	2453.1783	0.0054	0	100.42	54	9.40E-08	AGQNTYVQSNNSGIANVST SGTVGK
956.8341	2867.4803	3	2867.4777	0.0026	0	97.13	53	3.90E-07	VNFQDLISQQLSGAARPTE ELGGVNPK
1018.4899	3052.448	3	3052.441	0.007	0	87.63	53	3.40E-07	GYFIGGTLEMSNVDLTDQF VDMIVTQK[a]
905.8442	4524.1845	5	4524.1794	0.0052	0	41.45	54	1.70E-03	FDNNGHLASVTDTAGNVTS PAGQVLVQISYNVVGANPD EAGAPTR[b]

[a] - Methionine oxidation, [b] - NQ deamidation, [c] – 2 oxidation. **MZ** - Measured m/z ratio; **MR** – Average charge-adjusted mass; **Missed Sites** – uncleaved trypsin sites in peptide; **Mascot Score** – Statistical measure of sequence match based on MS/MS parent and fragment ion patterns; **Score Threshold** – Threshold above which Mascort Score will give a match P > 0.05 considering all peptides in Mascot data base. **Expect Value** – expectation value – number of matches with equal or better scores that would be expected to occur by chance.

Supplementary Table 5. *Borrelia burgdorferi* PH Unique Peptides. Native peptides identified from MS/MS of Bb polyhooks. Sequence coverage = 86.9%.

Measured MZ	Measured MR	Z	Calculated MR	ΔMR	Missed Sites	Mascot Score	Score Threshold	Expect Value	Sequence
526.7359	1051.4572	2	1051.4571	0.0002	0	66.41	46	5.30E-04	AGAFDVDSDR
526.7361	1051.4576	2	1051.4571	0.0005	0	69.89	47	2.70E-04	AGAFDVDSDR
714.6797	2141.0172	3	2141.0072	0.01	1	48.82	52	9.00E-04	AGAFDVDSDRHLVNPANGMR[b,c]
936.4437	3741.7455	4	3741.729	0.0165	1	72.39	52	7.80E-06	AIIQDGYGMGYMENYEIDQNGVI VGIYSNGIRR[a,b]
762.9066	1523.7987	2	1523.7984	0.0003	0	97.27	51	6.70E-07	ASDLGVSGNGFFILK
866.4569	2596.3489	3	2596.3483	0.0006	1	50.53	53	6.30E-04	DLEGEKVINTASDIEDLIIPIGDK
994.8557	2981.5453	3	2981.5444	0.0009	2	76.06	52	1.00E-05	DLEGEKVINTASDIEDLIIPIGDKE GAK
601.6661	1801.9764	3	1801.976	0.0004	1	47.02	52	4.30E-03	DLGKIALASFMNPGGLAK
1022.8871	3065.6394	3	3065.6397	-0.0003	0	74.91	50	8.70E-06	GDILQIPITFNVLGANVGEVGEQ QTVNLK
778.4207	2332.2403	3	2332.2387	0.0017	1	50.33	51	8.40E-04	GFQANAKTITTSDQLLQELVR
801.1532	3200.5837	4	3200.5885	-0.0048	1	39.71	54	2.90E-03	GGTNPKQVGLGMNVASIDTIHTQ GAFQSTQK[a]
624.0628	2492.2221	4	2492.219	0.0031	1	79.28	54	3.60E-06	GRVNFQDMISQSISGASRPTDA R
554.7871	1107.5597	2	1107.5607	-0.0011	0	45.74	51	2.10E-02	HLVNPANGMR
695.382	1388.7494	2	1388.7486	0.0008	0	77.62	52	4.90E-05	IALASFMNPGGLAK
1023.5078	3067.5016	3	3067.5033	-0.0017	1	69.59	54	4.70E-06	IALASFMNPGGLAKSGDTNFVET SNSGQVR
636.3503	1270.686	2	1270.6881	-0.0021	0	95.91	53	3.00E-06	IGETGLAGLGDIR
1083.0158	2164.0169	2	2164.0172	-0.0002	0	118.72	52	1.50E-10	LGTVGSYTDSITQFADSSSTK
789.4359	1576.8573	2	1576.8573	0	0	96.7	51	1.10E-07	LPLIQEGANPADIAR
903.4657	1804.9168	2	1804.9142	0.0027	0	83.43	53	5.20E-05	MDVVGNNIANVNTIGFK
967.5124	1933.0103	2	1933.0091	0.0012	1	81.64	53	2.10E-06	MDVVGNNIANVNTIGFKK
877.7788	2630.3146	3	2630.3123	0.0023	0	71.15	54	9.80E-04	QVGLGMNVASIDTIHTQGAFQST QK
578.6603	1732.959	3	1732.9584	0.0006	1	83.73	50	1.90E-05	RLPLIQEGANPADIAR
849.3904	1696.7662	2	1696.7653	0.0009	0	103.08	50	2.50E-07	SGDTNFVETSNSGQVR
984.1555	2949.4446	3	2949.4428	0.0017	1	82.03	54	2.90E-07	SGDTNFVETSNSGQVRIGETGL AGLGDIR
849.9389	1697.8632	2	1697.8624	0.0008	0	123.28	52	8.70E-10	SLYDSFGNVSVLELR
808.9464	1615.8781	2	1615.8781	0.0001	0	123.29	51	3.70E-09	TITTSDQLLQELVR
963.5253	1925.036	2	1925.0357	0.0003	0	84.85	51	1.90E-06	VINTASDIEDLIIPIGDK
771.0849	2310.2329	3	2310.2318	0.001	1	62.84	51	1.40E-04	VINTASDIEDLIIPIGDKEGAK
657.6086	2626.4055	4	2626.4065	-0.001	2	58.29	51	3.10E-04	VINTASDIEDLIIPIGDKEGAKSTK
760.7064	2279.0973	3	2279.0964	0.0008	0	63.98	53	1.40E-04	VNFQDMISQSISGASRPTDAR

[a] - Methionine oxidation, [b] - NQ deamidation, [c] - 2-oxidation. **MZ** - Measured m/z ratio; **MR** – Average charge-adjusted mass; **Missed Sites** – uncleaved trypsin sites in peptide; **Mascot Score** – Statistical measure of sequence match based on MS/MS parent and fragment ion patterns; **Score Threshold** – Threshold above which Mascort Score will give a match P > 0.05 considering all peptides in Mascot data base. **Expect Value** – expectation value – number of matches with equal or better scores that would be expected to occur by chance.

Supplementary Table 6. Oligonucleotide primers used in this study.

Primers	Sequences [a]	Note [b]
P_1	CACCATGATGAGATCATTATTTTCGGG	Over-expression of *TDE2768*; F
P_2	CTATCGTTTCAAGTTCAAGAC	Over-expression of *TDE2768*; R
P_3	ATGATGAGATCATTATTTTCGGG	Amplification of *TDE2768*; F
P_4	CCTATAGGTCAAGCACTTGATGCAAAGG	Site-mutagenesis of *TDE2768(165K-A)*; F
P_5	CCTTTGCATCAAGTGCTTGACCTATAGG	Site-mutagenesis of *TDE2768(165K-A)*; R
P_6	GGTCAAAAACTTGATGCAGCGGCAACCAC	Site-mutagenesis of *TDE2768(169K-A)*; F
P_7	GTGGTTGCCGCTGCATCAAGTTTTTGACC	Site-mutagenesis of *TDE2768(169K-A)*; R
P_8	CCACAAGTGTAGCCTATGCTTGTAACCTTG	Site-mutagenesis of *TDE2768(175N-A)*; F
P_9	CAAGGTTACAAGCATAGGCTACACTTGTGG	Site-mutagenesis of *TDE2768(175N-A)*; R
P_{10}	GTGTAAACTATGCTGCTAACCTTGATAAGAGGCTG	Site-mutagenesis of *TDE2768(178C-A)*; F
P_{11}	CAGCCTCTTATCAAGGTTAGCAGCATAGTTTACAC	Site-mutagenesis of *TDE2768(178C-A)*; R
P_{12}	CTATGCTTGTGCCCTTGATAAGAGGCTGCC	Site-mutagenesis of *TDE2768(179N-A)*; R
P_{13}	GGCAGCCTCTTATCAAGGGCACAAGCATAG	Site-mutagenesis of *TDE2768(179N-A)*; F
P_{14}	GCTTGTAACCTTGCTAAGAGGCTGCCTG	Site-mutagenesis of *TDE2768(181D-A)*; R
P_{15}	CAGGCAGCCTCTTAGCAAGGTTACAAGC	Site-mutagenesis of *TDE2768(181D-A)*; F
P_{16}	GGGATGCTTCAGCAGAC	5’ -flank region of *TDE2768*; F
P_{17}	GCTGCTGCGTAACATAATTATTGCCTCCTAATTG	5’ -flank region of *TDE2768- aacC*; R
P_{18}	ATGTTACGCAGCAGCAACGATG	*aacC* cassette; F
P_{19}	TTAGGTGGCGGTACTTGGGTC	*aacC* cassette; R
P_{20}	GTACCGCCACCTAGGTATGGTATAATATAGGG	3’ -flank region of *TDE2768- aacC*; F
P_{21}	CGGCTTGAATTCCAAGTAC	3’ -flank region of *TDE2768*; R
P_{22}	ATGAACAAAAATATAAAATATTCTC	*ermB* cassette; F
P_{23}	TTATTTCCTCCCGTTAAATAATAG	*ermB* cassette; R
P_{24}	*CCTAGG*AGGCAATAATTATGATGAGATCATTATTTTC	TdFlgE* formation, *TDE2768*; F
P_{25}	*CCTAGG*TTCGTTTCAAGTTCAAGACTG	TdFlgE* formation, *TDE2768*; R
P_{26}	GAGAATATTTTATATTTTTGTTCATAATTATTGCCT CCTAATTG	5’ -flank region of *TDE2768-ermB*; R
P_{27}	CTATTATTTAACGGGAGGAAATAA*CCTAGG*TATGG TATAATATAGGG	3’ -flank region of *TDE2768-ermB*; F

[a] Underlined portions show the engineered overlapping base pairs and italic show the AvrII restriction enzyme cutting-site; [b] Primer orientation: F, forward; R, reverse.

Videos: All videos were taken at 400X by dark-field microscopy.

Supplementary video 1. Cells WT *T. denticola* in 1% methylcelluose. Note cells had notable translational motility. Cells of replacement mutant TdFlgE* had identical swim behavior.

Supplementary video 2. Cells of mutant ΔflgE in 1% methylcellulose. Note cells deleted of *flgE* were completely non-motile.

Supplementary video 3. Cells of substitution mutant TdC178A in 1% methylcellulose. Note cells of TdC178A generated motion but lacked translational motility. Cells of substitution mutants TdK165A and TdN179A had identical swim behavior.

Supplementary video 4. Cells of substitution mutant TdK169A in 1% methylcellulose. Note cells of TdK169A had notable translational motility similar to that of the WT and TdFlgE*. Cells of substitution mutant TdD181A had identical swim behavior.