HHS Public Access

Author manuscript

Nat Chem Biol. Author manuscript; available in PMC 2020 February 12.

Published in final edited form as:

Nat Chem Biol. 2019 October ; 15(10): 959–965. doi:10.1038/s41589-019-0341-3.

Structure and chemistry of lysinoalanine crosslinking in the spirochaete flagella hook

Michael J. Lynch[1], **Michael Miller**[2], **Milinda James**[3], **Sheng Zhang**[4], **Kai Zhang**[5], **Chunhao Li**[5], **Nyles W. Charon**[3], **Brian R. Crane**[1,*]

[1]Department of Chemistry and Chemical Biology, Cornell University, Ithaca, NY, USA

[2]Department of Biochemistry, Robert C. Byrd Health Sciences Center, West Virginia University, Morgantown, WV, USA

[3]Department of Microbiology, Immunology, and Cell Biology, Robert C. Byrd Health Sciences Center, West Virginia University, Morgantown, WV, USA

[4]Proteomics Facility, Institute of Biotechnology, Cornell University, Ithaca, NY, USA

[5]Philips Institute for Oral Health Research, Virginia Commonwealth University School of Dentistry, Richmond, VA, USA

Abstract

The flagellar hook protein FlgE from spirochaete bacteria self-catalyzes the formation of an unusual inter-subunit lysinoalanine (Lal) crosslink that is critical for cell motility. Unlike other known examples of Lal biosynthesis, conserved cysteine and lysine residues in FlgE spontaneously react to form Lal without the involvement of additional enzymes. Oligomerization of FlgE via its D0 and Dc domains drives assembly of the crosslinking site at the D1–D2 domain interface. Structures of the $FlgE_{D2}$ domain, dehydroalanine (DHA) intermediate, and Lal crosslinked FlgE subunits reveal successive snapshots of the reaction. Cys178 flips from a buried configuration to release hydrogen sulfide (H_2S/HS^-) and produce DHA. Interface residues provide hydrogen bonds to anchor the active site, facilitate β-elimination of Cys178, and polarize the peptide backbone to activate DHA for reaction with Lys165. Cysteine-reactive molecules accelerate DHA formation, whereas nucleophiles can intercept the DHA intermediate, thereby indicating a potential for Lal crosslink inhibitors to combat spirochaetal diseases.

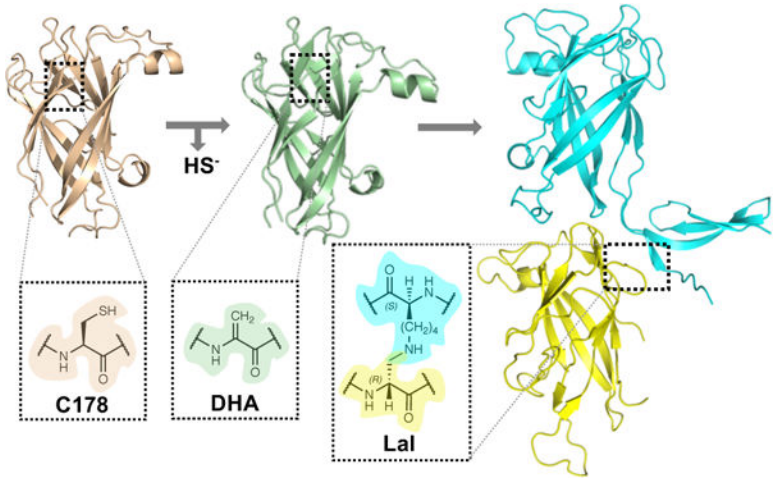
Graphical Abstract

*Corresponding author: bc69@cornell.edu, phone: 607-254-8634.

Author Contributions:

N.W.C., B.R.C., C.L., and M.M. conceived the study. K.Z. generated an initial construct encoding for Td FlgE. M.J.L. performed the crystallography, X-ray diffraction experiments and associated structural analysis, site-directed mutagenesis, EDC crosslinking and Lal crosslinking SDS-PAGE assays. M.J. carried out preliminary work that identified NEM and DTNB as FlgE crosslink enhancers. M.J.L. prepared samples for MS. S.Z. performed MS data acquisition, associated samples analysis and prepared the MS peptide quantification methods section. M.J.L. and B.R.C. prepared figures. M.J.L., B.R.C., and N.W.C wrote the manuscript. B.R.C., N.W.C., M.J.L., C.L., M.J., M.M., and S.Z helped edit the manuscript. B.R.C., N.W.C., M.M. and C.L. supervised the study.

Competing Interest: The authors declare no competing interests.



Introduction:

Genetically encoded protein crosslinks are rare in biology, with only four general categories known (disulfide, isopeptide, ester and thioester bonds).[1–6] Recently, we have discovered a new type of inter-molecular protein crosslink in the form of lysinoalanine (Lal), which forms spontaneously between conserved cysteine and lysine residues in the flagellar hook protein FlgE of spirochaetes.[7] Unlike most flagellated bacteria, spirochaetes enclose their hooks and filaments in the periplasm (Fig. 1a).[8,9] The ion gradient spanning the inner membrane drives the rotation of the periplasmic flagella (PFs) and propels the cell forward in a characteristic corkscrew-like or beating wave trajectory.[8] The hook is a 55 +/− 7 nm long structure[8,10,11] that connects the PF to the motor embedded in the cell body and serves as a universal joint to smoothly transmit motor torque to the PFs.[12] As the hook rotates, its constituent ~120–130 FlgE subunits[13] dramatically alter their interactions in response to the large changes in hook curvature. In spirochaetes, hooks must operate in the confined environment of the periplasm,[14] where the PFs interact with the cell cylinder to generate backward moving waves along the cell body. Thus, compared to most motile bacteria with external flagella, a greater need for hook stabilization and fortification in spirochaetes likely underlies Lal crosslinking in the FlgE subunits.

Two examples indicate that covalent crosslinking adds strength to macromolecular structures. First, pili and cell-surface adhesins from gram-positive bacteria have protein crosslinks in the form of ester and isopeptide bonds.[4] Mutant pili proteins that are deficient in one type of crosslinks are considerably more extensible than those of the wild type.[15] Furthermore, *Escherichia coli* (Ec) phage HK97 have head proteins that are extensively crosslinked to one another via inter-subunit isopeptide bonds.[5,16] The stability of wild type capsids far exceeds that of mutants that lack these crosslinks.[17,18]

Although Lal is not previously known to strengthen protein assemblies, it is found in the preparation of protein-containing foods that have been treated with heat or alkaline solutions.[19] In these cases, Lal formation involves a multi-step reaction initiated by the base-catalyzed deprotonation of a cysteine or serine α-proton by a hydroxide ion, followed by the

non-specific addition of a lysine side chain ε-amine to the alkene intermediate.[19] Conversely, natural Lal production in biological systems is strictly controlled, and limited to a group of ribosomally synthesized and post-translationally modified peptides known as lanthipeptides.[20,21] Most lanthipeptides contain the thioether cross-linkage lanthionine and/or methyllanthionine; however, duramycin, duramycin B, C, and cinnamycin represent a unique subset of lanthipeptides that also contain a Lal crosslink.[21,22] Recently, isolation and characterization of the duramycin biosynthetic gene cluster from *Streptomyces cinnamoneus* revealed that as many as three different enzymes are required for Lal biosynthesis, with the terminal enzyme DurN catalyzing a substrate-assisted addition in order to form Lal.[21,23]

In contrast to Lal biogenesis in lanthipeptides, Lal biosynthesis by spirocheate FlgE is auto-catalytic and requires no other enzymes or cofactors.[7] In the oral pathogen *Treponema denticola* (Td), FlgE is a 50 kDa protein composed of four distinct domains: D0, Dc, D1, and D2 (Fig. 1b). The Lal crosslink forms between Cys178 of D2 and Lys165 of the D1–D2 linker in the high-molecular-weight complex (HMWC) isolated from polyhook structures in strains of Td, the Lyme disease spirochaete *Borrelia burgdoferi*[7,24], and polymerized recombinant Td FlgE.[7] In addition to C178 and K165, Lal crosslinking depends on the invariant residue Asn179 (ref. [7]) and proceeds through a dehydroalanine (DHA) intermediate (Fig 1c). In an interesting parallel, piperidinyl-alanine formation is a common side reaction in the solid-phase synthesis of cysteine-containing peptides that also involves conversion of a sulfhydryl-modified side chain to DHA and Michael addition of an amine.[25] Despite some knowledge of the key residues and intermediates involved in FlgE Lal biosynthesis,[7] the underlying mechanism and structural constraints on that mechanism are largely unknown.

Herein we reveal the structural and chemical requirements of Lal crosslinking by Td FlgE. Spirochaetes are known for widespread diseases such as Lyme, syphilis, leptospirosis and periodontal disease.[7,9] Lal crosslink-deficient mutants of Td FlgE display normal hook assemblies and cellular morphology, but exhibit defects in motility,[7] thereby revealing the potential antimicrobial utility of Lal crosslink inhibitors in combating spirochaete-related human diseases. Defining the key features of Lal formation by FlgE also enables its use in the protein engineering of covalently-linked oligomers that are resistant to hydrolysis and redox modifications, a benefit for drug development.

Results:

FlgE oligomerization precedes Lal crosslink formation

In Td FlgE, the D0 domains reside on the N and C termini (Fig. 1b) and share high sequence homology to the D0 domains of other axial flagellar motor proteins (e.g. the rod and filament subunits).[26,27] Previously, in-frame N- and C-terminal D0 and Dc domain deletions of *Salmonella* FlgE were shown to interfere with hook assembly, suggesting that inter-subunit D0–Dc domain interactions are essential for proper hook formation and function.[28,29] Indeed, cryo-EM structures of fully assembled flagellar hooks from *Salmonella enterica* and *Campylobacter jejuni* show extensive inter-subunit domain interactions mediated by the D0 and Dc domains.[30–32] Based on these observations, we assessed the dependence of *in vitro* Lal crosslinking on the presence of the Td FlgE D0 and Dc domains.

Crosslinked FlgE HMWCs readily form from purified full-length FlgE ($FlgE_{FL}$) monomer and this reaction requires the D0 and Dc domains (Fig. 2a, Supplementary Fig. 1a). FlgE oligomerization, captured with the heterobifunctional protein crosslinking reagent 1-ethyl-3-(3-dimethylaminopropyl)carbodiimide hydrochloride (EDC)[33] occurs prior to Lal formation and does not depend on the active site residues K165, C178, or N179 that have been previously shown to be required for Lal crosslinking (Supplementary Fig. 1b,c and Fig. 2b).[7] These data suggest that initial FlgE oligomerization depends on the D0 and Dc domains and occurs faster than Lal crosslinking.

Interestingly, treatment of wild type (WT, C178) $FlgE_{FL}$ with the cysteine-reactive compounds N-ethylmaleimide (NEM) or 5,5-dithio-bis-(2-nitrobenzoic acid) (DTNB) accelerates, rather than blocks, HMWC formation in WT $FlgE_{FL}$ and rescues HMWC formation in crosslinking deficient FlgE D0/Dc domain truncations (Supplementary Fig. 1d). HPLC-MS analysis of $FlgE_{T31\text{-}M434}$ treated with DTNB produced 2-nitro-5-(sulfinothio)benzoic acid ($[HSO_2\text{-TNB}]^{-1}$, m/z = 261.95, Supplementary Fig. 2), thereby indicating that DTNB (and presumably NEM) aids the elimination of the C178 thiol by converting it to a better leaving group. On the other hand, β-mercaptoethanol (βME) blocks crosslinking by competing with K165, intercepting the DHA intermediate and forming a thioether adduct (Fig 2c, d). Thus, Cys alkylation or disulfide formation facilitates elimination to DHA, which is then reactive toward nucleophiles.

C178 to DHA conversion releases hydrogen sulfide

In lanthipeptides, serine and threonine residues are common precursors of β-elimination reactions; however, these residues often require activation via either phosphorylation or glutamylation of the side-chain hydroxyl group.[20,34] In contrast, elimination of an unmodified cysteine thiol is rare. Notably, *Leptospira* FlgE is the only spirochaete hook protein that has a residue other than cysteine, in this case serine, at the Lal crosslinking position. In Td FlgE, substitution of C178 to serine, but not threonine, results in a variant that still crosslinks, albeit at reduced levels and with a different pH dependence compared to WT (Fig. 2b, Supplementary Fig. 3). Unsuccessful efforts to trap alternative C178 oxoforms or perturb FlgE crosslinking with hydrogen peroxide or metal ions indicate that Lal crosslinking does not depend on C178 oxidation or thiol metallo-coordination (Supplementary Fig. 4a,b), leaving the alternative that the cysteinyl thiol eliminates directly as hydrogen sulfide (H_2S/HS^-). Sulfide measurements with the NBD-N_3 sulfide-specific spectrophotometric probe[35] produced 35 +/− 5 μM sulfide from 60 μM FlgE, equating to 65 +/− 8% of the total FlgE concentration (Fig. 2e and Supplementary Fig. 5a–c). This quantity correlates well with the 59% HMWC formation observed on SDS-PAGE (Fig. 2a) and was reproduced by an orthogonal sulfide detection assay that we developed specifically for the Td FlgE Lal crosslinking system.[36] This latter assay utilizes the *E. coli* O-acetylserine sulfhydrylase CysM to convert released sulfide to the more stable L-cysteine product *in situ* and thereby allows direct capture of the unmodified H_2S/HS^- product in the presence of ambient oxidants (Supplementary Fig. 5d–f).[36]

Structural characterization of WT, DHA, and Lal crosslinked FlgE

To better understand the chemical environment that drives Lal formation, we determined the structures of WT, DHA, and Lal crosslinked Td FlgE (Fig. 3a–b, Supplementary Table 1). Similar to the other FlgE proteins[32,37,38], the Td $FlgE_{D2}$ domain folds into a jelly-roll motif comprised of two β-sheets folded together and stabilized by two α-helical linkers and coil regions (Fig. 3a and Supplementary Fig. 6). C178 resides in the first β-strand of $FlgE_{D2}$ where it projects inwards toward the center of the β-barrel and is surrounded by Y176, W201, and K336, along with N179 (Fig. 3a, top inset). The DHA intermediate is fleeting in the course of normal Lal crosslinking; thus, to produce the intermediate in high yield, WT Td $FlgE_{D2}$ was treated with the cysteine alkylating agent 2,5-dibromohexanediamide (DBA)[39,40] and DHA formation was confirmed by LC-MS (Supplementary Fig. 7a). Although the global structure of the DHA $FlgE_{D2}$ domain remains unperturbed compared to WT (RMSD = 0.472), C178 conversion results in substantial local conformation changes that include the melting of the β-strand harboring residue 178 (green, Fig. 3a, bottom inset). Unbiased electron-density maps reveal the loss of the cysteinyl thiol electron density, an increased planarity about the DHA Cα carbon and a shortening of the Cα–Cβ bond distance (Fig. 3c). The planarity of DHA provides increased accessibility of the Cβ Michael acceptor for nucleophilic attack by K165 from an adjacent FlgE subunit (Fig. 3d). There are no basic residues in the vicinity of DHA to deprotonate the C178 Hα precursor; however, two structured solvent molecules coordinated by S322, T325, T327, and the peptide amide of DHA are within 3.8 Å of the DHA Cα (Supplementary Fig. 8a,b). Modeling a cysteine residue into the DHA configuration places these solvent molecules within ~2.4 Å from the C178 Hα (Supplementary Fig. 8c). Considering that the pKa difference between the C178 Hα and hydronium ion is large (likely > 15) and Td FlgE Lal crosslinking assays are carried out in slightly alkaline buffer conditions (pH 8.5), a hydroxide ion may deprotonate C178 Hα. Alternatively, because the conformation of Cys178 during β-elimination is unknown, another possibility is that the thiol moiety itself deprontonates Hα[41] or activates ordered solvent to do so.

DHA formation also results in a cascade of conformational and hydrogen bonding perturbations between key residues in the vicinity of position 178 (Y176, W201, F205, T334, T335, K336, and Y338; Supplementary Fig. 9a–h). Notably, the position of N179 remains unchanged in all three structures (Supplementary Fig. 9a,i), despite the N179A variant forming no Lal crosslinked HMWCs (Fig. 2b) and causing non-translating cells.[7] We considered that the N179 sidechain may participate in the elimination reaction through formation of a thioether or thioamide with the C178 thiol. However, an $^{18}OH_2$ substitution experiment showed no evidence of ^{18}O incorporation anywhere within the crosslinked peptide, which would be expected upon resolution of a thiol adduct formed between C178 and N179, or another residue. Structurally, the N179 amide sidechain stabilizes the β-strand harboring the C178/DHA residue by simultaneously hydrogen bonding to two main-chain peptide bonds across a broad turn that encompass the C178/DHA active site. Unlike N179, several other surrounding residues undergo large changes in sidechain conformation and hydrogen bonding partners, including K336, four threonine residues (T325, T327, T334, and T335) and a chain of aromatic residues (Y176, F205, Y338, and W201, Supplementary Fig. 9b–h). Mapping the electrostatic potential onto the molecular surface of WT and DHA

$FlgE_{D2}$ reveals an overall remodeling of the $FlgE_{D2}$ surface in the vicinity of the active site coupled to a decrease in the net positivity of the DHA pocket compared to WT (Supplementary Fig. 10a). Neutralization of the DHA active site pocket in this manner may potentially promote Lal formation by favoring deprotonation of the K165 side chain.

Crystallization of crosslinked FlgE subunits was prevented by heterogeneity of the recombinant polymers generated during *in vitro* Td FlgE crosslinking. Thus, to determine the Lal crosslink structure, we reacted DHA $FlgE_{D2}$ with a crosslinking-deficient $FlgE_{D1D2}$ domain truncation ($FlgE_{N91\text{-}S432}$ C178A). When mixed with $FlgE_{N91\text{-}S432}$ C178A, DHA $FlgE_{D2,}$ but not WT $FlgE_{D2}$, formed a dimer with the expected molecular weight of 54 kDa (Supplementary Fig. 7b). Thus, C178 conversion to DHA activates FlgE for crosslinking. Optimization, large-scale purification (Supplementary Fig. 7c–e), and sparse-course matrix screening afforded diffraction-quality crystals which were used to determine the Lal crosslink structure contained within the $FlgE_{D1D2:D2}$ dimer (Fig 3b). Consistent with the DHA orientation and juxtaposition of the adjacent FlgE, the Lal crosslink contains D- (or *R*) rather than L- stereochemistry at residue 178 (Fig. 3b, inset). The crosslink structure reveals two ordered solvent molecules coordinated within 3 Å of Lal in a similar mode to that found in the DHA structure (Supplementary Fig. 8d–f). There are no basic residues supplied by either subunit to the interface (Supplementary Fig. 9h–j) for deprotonation of the K165 ε-amine prior to Michael addition. Analogous to its proposed role in the β-elimination of C178, hydroxide or activated water may again serve as a proton acceptor to activate K165 for nucleophilic attack. Following K165 attack on DHA, the *Si* face of the resulting enolate faces solvent for reprotonation. The Lal crosslink structure also highlights the role of complimentary surface charge and hydrogen-bonding interactions for stabilizing the $FlgE_{D1D2:D2}$ interface (Supplementary Fig. 10b–d).

Subunit association in the Lal crosslinked structure agrees well with low-resolution cryoEM maps of the flagellar hook from *S. enterica* (Fig. 3e).[42] The Td FlgE Lal crosslinked dimer is well accommodated in whole-hook electron density without adjustment of the individual FlgE subunit protomers. Furthermore, the small region consisting of the Td FlgE D1 domain projects inward toward the center of the hook and coincides with a lobe of density that was originally assigned to the D1 domain of *S. enterica* FlgE.[42]

Residue contributions to Td FlgE crosslinking

In order to assess the importance of specific residues identified by the structural analysis of WT, DHA, and Lal crosslinked FlgE, we carried out structure-guided site-directed mutagenesis and monitored the crosslinking of each variant. Substitution of conserved FlgE residues highlight stringent constraints on the nucleophilicity of the 165 position (K165H, K165R), the spatial positioning of the DHA precursor (Asn-Cys, and Cys-Ala sequence permutations) and structural stabilization by Asn179 (N179D/E/Q), as well as identify another essential residue for crosslinking, T334 (Fig. 2b, Supplementary Fig. 11a,b). Furthermore, altering residues that contact C178 within the β-barrel (Y176, W201, K336; Fig. 3a) does not dramatically inhibit HMWC formation (Fig. 2b), supporting that implication from C178 reactivity toward NEM and DTNB that β-elimination follows C178 repositioning to the surface. The amount of DHA produced prior to crosslinking is also not

greatly affected by the K165A substitution (Fig. 2d). It follows that K165 is unlikely to act as an essential base in the deprotonation of the C178 Hα.

FlgE residue substitutions do affect the ability of the inhibitor βME to intercept the DHA intermediate and act as a surrogate nucleophile for K165. In particular, substitutions of N179 and T334 prevent the Michael addition step of Lal biogenesis (Fig. 2d) and also reaction with βME, for which there are less orientational restraints. N179A likely disrupts the β-turn that positions DHA and its associated hydrogen bond network. In contrast, T334 appears less important for structural stabilization and may rather activate the Michael acceptor. Michael acceptors are deactivated in DHA-peptide mimics such as acrylamide because of electron donation of the nitrogen lone pair into the ene-one moiety.[43] Delocalization of the nitrogen lone pair away from the neighboring carbonyl group by interaction with protic solvents or an adjacent hydrogen bond donating group, for example, mitigates this effect.[43] T334 likely serves a similar role in FlgE by donating a hydrogen bond to the N179 peptide nitrogen to disfavor the imine resonance form of the peptide bond and thereby favor the ketone and enolate forms of the Michael acceptor (Fig. 4). Polarization of the DHA ene-one toward the more reactive enolate is further enhanced by main-chain hydrogen bonds to the A177 and N179 carbonyl groups from K336 and T327. Substitution of either residue to Pro, but not Ala, blocks crosslinking (Fig. 2b). Notably, in one copy of the two Lal crosslinked dimers present in the until cell, the hydrogen bond between the T334 sidechain and the N179 peptide nitrogen decreases substantially from ~2.9 Å in DHA $FlgE_{D2}$ to ~2.4 Å. Contraction of this hydrogen bond suggests a switch of T334 from a hydrogen bond donor to an acceptor, as well as an increase in imine character of the Lal178-N179 peptide bond (Fig. 4). Interestingly, the amount of the BME-adduct peptide produced by the K165A variant is also reduced relative to WT Td $FlgE_{D2}$ (Fig. 2d). The role of K165 in βME-adduct formation is unclear but may involve interactions with the peptide backbone, ordered solvent or βME itself.

Discussion:

Prior to FlgE, specific biosynthesis of Lal was limited to lanthipeptides, a class of ribosomally derived post-translationally modified peptides.[20,21] For Lal formation within the lanthipeptide duramycin by DurN, an additional enzyme DurX generates an essential cofactor – erythro-3-hydroxy-L-aspartic acid (Hya) – to assist in Lal formation.[23] Hya acts as a general base to deprotonate the Lal precursor lysine ε-amine for nucleophilic attack on DHA, which is previously formed via serine phosphorylation and elimination.[23]

In contrast to Lal formation in duramycin, a very different biochemical mechanism produces Lal in the spirochaete flagellar hook. Through our structural and biochemical experiments, we show that an active center composed from the FlgE subunit interface is sufficient to self-catalyze formation of the unusual Lal crosslink, and does so without the need for additional substrates, enzymes, or cofactors. During natural hook assembly, FlgE is secreted through the central pore of the growing flagellar rod before polymerizing into the newly forming hook.[44,45] Thus, it is not surprisingly that the progression of *in vitro* Lal crosslinking first involves the assembly of $FlgE_{FL}$ into non-covalent oligomers mediated by D0–Dc domain inter-subunit interactions (Fig. 4a and Supplementary Fig. 1). C178 is then converted to

DHA through the β-elimination of the cysteine thiol (Fig. 4b), releasing hydrogen sulfide as a byproduct (Fig. 4b and Supplementary Fig. 5).[36] C178 exposure likely precedes hydrogen sulfide release since C178 readily reacts with compounds (NEM and DTNB, Supplementary Fig. 2) that are too large to access the buried pocket of the D2 domain β-barrel, and changes to C178 interacting residues in this pocket do not greatly inhibit crosslinking and hence β-elimination (Fig. 2b).

An interesting issue is the identity of the base required to deprotonate C178 Cα for release of unmodified sulfide. Because DHA formation does not depend on the crosslinking residue K165 and there are no other essential basic residues in the interface, coordinated solvent molecules found in the $FlgE_{D2}$ DHA structure may serve as proton acceptors for the C178 Hα (B in Fig. 4b). To have the strength to extract the C178 Cα proton, these solvent molecules may take the form of hydroxide or undergo a Grotthuss-type proton transfer with another base, which could be the exposed C178 thiol itself.[41] It is also possible that conformational stabilization of the planar DHA conformation by the protein increases the acidity of the C178 Hα. Likewise, with the exception of ordered solvent, there is no adjacent base to deprotonate the K165 ε-amine, although the reduced electrostatic character of the interface may favor the neutral form. We only observe small amounts of the DHA intermediate (< 2 %) and it does not accumulate in any of the variants; thus, the two steps of the reaction may be quite concerted. Under such a regime, both the C178 thiol and the K165 amine may contribute to proton management, along with solvent.

Regardless of the proton acceptors, both steps of the reaction are likely driven by hydrogen bond networks that stabilize the DHA ene-one product and then activate it for Michael addition (Fig. 4c). The highly deleterious effects of any substitution at position 179 underscore the importance of a structural scaffold to Lal crosslinking. DHA formation is rate-limiting, but when produced in high yield, DHA readily promotes lysine addition, as evidenced by the ability of DHA-containing $FlgE_{D2}$ to readily crosslink with the weakly associating $FlgE_{N91\text{-}S423}$ (WT or C178A) truncation variant (Supplementary Fig. 7b). For the reaction with K165, key hydrogen bonds between the amide nitrogen of N179 and sidechain hydroxyl of T334 activate DHA (Fig. 4c). By engaging with the amide nitrogen lone pair of N179, T334 prevents lone pair delocalization into the peptide backbone, thereby polarizing the Cα-Cβ bond of DHA and increasing the overall electrophilicity of the DHA Michael acceptor. Upon successful installation of Lal, this hydrogen bond shrinks from 2.9 Å to 2.4 Å, indicative of a potential switch of T334 from a hydrogen bond donor to acceptor and increased contribution from the imine-containing resonance form of the Lal178-N179 peptide bond (Fig. 4c). Importantly, owing to the juxtaposition of the $FlgE_{D2}$ DHA residue relative to the K165 sidechain position in the interface, and to solvent exposure of the DHA *Si* face, Td FlgE extends stereochemical control over the reaction to form a Lal crosslink composed of L- or *(S)* K165 and D- or *(R)* A178 (Fig. 4c).

Overall, an increased understanding of how FlgE self-catalyzes the formation of the Lal crosslink empowers the development of novel motility-targeted antibiotic agents specific to spirochaete-related illnesses. Development of new antimicrobials to spirochaetes is more important now than ever, due in part to the increasing resistance of spirochaetes to front line drugs.[46,47] Similar to the often-targeted chemistry that links peptidoglycan in the bacterial

cell wall, FlgE Lal formation, also in the periplasm, has no known parallel in mammals and thus, crosslinking inhibitors are well positioned to avoid cross-reactivity with host metabolism.[48] Finally, DHA is increasingly being used for protein engineering and proteomic studies;[39,40,49,50] therefore the understanding of its formation and reactivity in the context of FlgE may advance its utility for such applications.

Online Methods:

- FlgE cloning and site-directed mutagenesis

FlgE was PCR cloned from *Treponema denticola* ATCC 35405 genomic DNA and inserted into pet28a between NdeI and BamHI restriction sites. Site-directed mutagenesis was carried out using the QuikChange method and confirmed by DNA sequencing.

- Protein expression, refolding, and purification

FlgE variants were transformed into BL21 (DE3) *E. coli* cells, grown in LB media containing 50 μg/mL kanamycin to an OD_{600} of ~ 0.6–0.8 at 37 °C and induced with 1 mM IPTG at 25 °C. Cells were harvested after 12–18 hours by centrifugation at 5,000 rpm at 4 °C. Full-length Td FlgE ($FlgE_{FL}$, WT and mutant constructs) was purified under denaturing conditions due to its tendency to form inclusion bodies. Briefly, cell pellets were re-suspended in 60 mL of lysis buffer (50 mM Tris pH 7.5, 500 mM NaCl, 10 mM imidazole, 6 M urea) and sonicated on ice. Lysates were centrifuged at 20,000 rpm for 45 minutes at 4 °C and purified by Ni-NTA affinity chromatography. $FlgE_{FL}$ was refolded via step-wise dialysis for 48 hours at 4°C. Briefly, 7.5 mL of Ni-NTA elution was rapidly diluted to 15 mL with 18MΩ H_2O and dialyzed into 1 L of 50 mM Tris pH 7.5, 150 mM NaCl, 5% (w/v) sucrose, 100 mM arginine hydrochloride, 1 M urea for 24 hours. $FlgE_{FL}$ was then transferred to 2 L of 50 mM Tris pH 7.5, 150 mM NaCl and dialyzed for 24 hours. Aggregated protein was pelleted by centrifugation at 3,000 rpm and the supernatant passed over a size-exclusion column (SEC 26/60, Superdex-200, 17104301, GE Life Sciences) in 50 mM Tris pH 7.5, 150 mM NaCl. All truncated forms of TdFlgE expressed in the soluble fraction and were purified under non-denaturing conditions.

- SDS-PAGE lysinoalanine crosslinking assays

WT and mutant $FlgE_{FL}$ proteins were prepared in crosslinking buffer (XLB, 40 mM Tris pH 8.5, 160 mM NaCl, 1 M $(NH_4)_2SO_4$) at a final concentration and volume of 10 μM and 20 μL, respectively, and incubated at 4°C for 48 hours. Stock solution of NEM (99%, 139675, Beantown Chemical) and DTNB (D-8130, Sigma) were prepared to a final concentration of 20 mM in 100% ethanol. NEM and DTNB treated samples were prepared in a similar manner to non-treated samples except for the addition of 1.0 μL of 20 mM NEM/DTNB to yield a final concentration of 1 mM. Non-treated control samples were supplemented with 1.0 μL 100% ethanol. Samples were quenched with 6x SDS loading dye containing excess BME and electrophoresed on a 4–20% Tris-Gly SDS-PAGE gel. The relative crosslinking percentages were determined by comparing HMWC band intensities to that of the monomer band using the image analysis software FIJI.[51]

- Generation of the dehydroalanine form of $FlgE_{D2}$

To form $FlgE_{D2}$ DHA, 2,5-dibromohexanediamide (DBA) was synthesized as previously described.[18,19] Briefly, 12.5 g (85.6 mmol) adipic acid (470300–078, Ward's Science) was added to 37.5 mL (517 mmol) thionyl chloride (>99%, 230464, Sigma) and stirred under reflux at 80 °C for 1.5 hours. The solution was then cooled to 25 °C at which point 50 mL CCl_4 and 36.6 g (206 mmol) N-bromosuccinimide (99%, B81255, Sigma) was added with continuous stirring. To the reaction mixture, 10 drops of 8.9 M HBr was added slowly and heated to 80 °C under reflux for 2 hours. The reaction was then cooled to 0 °C to precipitate the acid chloride, which was then filtered, dried, and re-dissolved in 100% ethanol to afford a dark red liquid. In a 500-mL round bottom flask, 200 mL 14.8 M NH_4OH was cooled to 0 °C, and the acid chloride added dropwise with constant stirring. After all of the acid chloride addition was added, the solution was stirred for an additional 1 hour and the solid precipitate was collected and partially dried under vacuum. To purify the crude DBA product, the precipitate was mixed with 200 mL 50/50 H_2O:methanol for 30 minutes at 60 °C to afford a bright white solid that was filtered and dried overnight under vacuum. The presence and purity of the final DBA product was confirmed by ^1H NMR. To form DHA-enriched $FlgE_{D2}$, excess DBA was dissolved in DMSO and incubated with WT $FlgE_{D2}$ in 100 mM Tris pH 8.0 for one hour at 25°C, followed by four hours at 37 °C and twelve hours at 4 °C. $FlgE_{D2}$ DHA formation was confirmed by the detection of the tetrahydrothiophene leaving group via positive-ion LC/MS, and purified by analytical SEC (10/300, Superdex-75, 17104401, GE Life Sciences).

- LC-MS of dehydroalanine formation following treatment of $FlgE_{D2}$ with DBA

To confirm DHA $FLgE_{D2}$ formation after treatment with the thiol bis-alkylated agent DBA, LC-MS was performed. Briefly, 10 mg/mL (556 μM) WT $FlgE_{D2}$ in 100 mM Tris pH 8.0 was treated with excess DBA as described previously.[18,19] Protein was removed by centrifugation at 13,000 rpm at 4 °C for 20 minutes in a 10 kDa MWCO spin column. The flow through was then collected, flash frozen in liquid nitrogen, and stored at −80°C. For HPLC-MS, 1 μL of each sample (buffer + DBA, buffer + $FlgE_{D2}$, and $FlgE_{D2}$ + DBA) was injected into an HPLC Micromass Quattro Ultima Pt Mass Spectrometer (Agilent Eclipse XDB-C18 4.6 × 250mm 5.0um-particle size) pre-equilibrated with 99:1 H_2O:acetonitrile, both containing 0.1% acetic acid (v/v). Each sample was eluted at flow rate of 1.0 mL/min using the following gradient elution parameters (time (min), acetonitrile %): 00:00–05:00 (1%), 05:00–15:00 (1% to 100%), 15:00–20:00 (100%), 20:00–20:06 (100% to 1%), 20:06–25:00 (1%). The THT byproduct was detected in positive ion mode with a m/z of 175.0 +/−1 Da. All HPLC-MS data were manipulated and analyzed in Xcalibur operation software (Thermo-Fisher Scientific).

- $FlgE_{D1D2:D2}$ dimer formation

To form the $FlgE_{D1D2:D2}$ dimer, 10 μM $FlgE_{A168\text{-}T344}$ ($FlgE_{D2}$) was incubated with 10 μM $FlgE_{N91\text{-}S423}$ ($FlgE_{D1D2}$) C178A in XLB for 48 hours at 4 °C. To each 20 μL sample, 5 μL 6x SDS loading dye and 1 μL 14.3 M BME was added to denature and reduce samples prior to SDS-PAGE.

- WT $FlgE_{D2}$, DHA $FlgE_{D2}$, and LA crosslink $FlgE_{D1D2:D2}$ dimer crystallization and structure determination

WT $FlgE_{D2}$ was crystallized after 5 days at 17 °C in 100 mM HEPES pH 7.5, 2 M ammonium sulfate at a concentration of 5 mg/mL. DHA $FlgE_{D2}$ was crystallized in 1.6 M sodium citrate tribasic pH 8.2 in 4 days at 25 °C at a concentration of 12.5 mg/mL. Lal crosslink $FlgE_{D1D2:D2}$ dimer was crystallized in 200 mM magnesium acetate, 20% PEG 3350 (w/v) in 3 weeks at 25 °C. Additional $FlgE_{D1D2:D2}$ dimer crystals were formed after 1 month at 25 °C in 100 mM Tris-HCl pH 8.5, 20% PEG 8000 (w/v), 200 mM magnesium chloride. Data sets were integrated and scaled in HKL2000 and SCALEPACK.[52] The data sets were collected under the following parameters (structure – synchrotron, beamline, wavelength, and collection temperature): WT $FlgE_{D2}$ – CHESS, F1, 0.97700 Å, 77 K; DHA $FlgE_{D2}$ – APS, ID-C, 1,10500 Å, 77 K; and Lal crosslinked $FlgE_{D1D2:D2}$ – APS, ID-E, 0.97918 Å, 77 K. The structure of WT $FlgE_{D2}$ was determined by molecular replacement (MR) using a mixed model generated from the $FlgE_{D2}$ from *Salmonella* (PDB ID: 1WLG).[53] The DHA $FlgE_{D2}$ and Lal crosslink $FlgE_{D1D2:D2}$ structures were determined by MR using WT $FlgE_{D2}$. Manual adjustments were made with COOT and refinements made with Phenix Refine.[54–56] Statistics for data processing and refinement are reported in Supplementary Table 1. The final WT, DHA, and Lal crosslink structures had a R_{work}/R_{free} (%/%) of 15.8/17.9, 18.9/19.9 and 23.3/30.5. Unbiased omit maps were generated by 5 cycles of refinement in Phenix Refine with simulating annealing (cartesian and torsion). The Ramachandran and rotamer outliers (%) for each structure are as follows (structure - Ram. outliers/rotamer outliers): WT $FlgE_{D2}$ – 0.00/0.69, DHA $FlgE_{D2}$ – 0.00/1.40, and Lal crosslinked $FlgE_{D1D2:D2}$ – 6.60/0.16. Figures and bond distance measurements were made in COOT and PyMol.[54,56,57]

- EDC protein:protein crosslinking assay

Samples for EDC crosslinking were prepared in 50 mM HEPES pH 7.5, 150 mM NaCl at a concentration of 10 μM FlgE and incubated for 30 minutes at 4 °C. EDC (22980, Thermo) was added to yield a concentration of 20 mM and incubated at room temperature for 30 minutes. Crosslinking was quenched by the addition of 5 μL 1M Tris pH 7.5 and analyzed by SDS-PAGE.

- Measurement of sulfide production in $FlgE_{FL}$

Sulfide release was measured using the sulfide reactive probe NBD-N_3. NBD-N_3 was synthesized from sodium azide (19038–1000, Acros) and 4-chloro-7-nitro-1,2,3-benzoxadiazole (>97%, Lancaster) with no notable procedure changes and confirmed by ^1H NMR.[16] A 100 mM NBD-N_3 stock was freshly prepared the day of use in 100% acetonitrile and filtered through a 0.22 μm filter. Sodium sulfide calibration curve standards were prepared in triplicate to a final volume of 100 μL in XLB containing 1 mM NBD-N_3. After sulfide addition, samples were incubated anaerobically for 30 minutes, and their absorbance measured at 470 nm. The UV-Vis spectrophotometer was blanked with XLB containing 1 mM NBD-N_3. The standard samples were linear over a 0 – 60 μM sulfide range with a correlation coefficient of ~0.98. $FlgE_{FL}$ samples (WT and C178A) were prepared in quadruplet to a final concentration of 60 μM. Samples were incubated anaerobically for 48

hours at room temperature, centrifuged at 13,000 rpm for 10 minutes, and the absorbance measured at 470 nm (Agilent UV-Vis spectrophotometer 8453).

- Quantification of dehydroalanine and BME-adduct containing tryptic peptides

FlgE samples were resolved by a 10% SDS-PAGE. The FlgE protein bands from SDS-PAGE gels were cut into ~1 mm cubes and subjected to in-gel digestion with trypsin (Promega) followed by extraction of the tryptic peptide as reported previously.[58] The excised gel pieces were washed consecutively in 200 μL distilled water, 100 mM ammonium bicarbonate (Ambic)/ACN (1:1) and ACN. The gel pieces were directly alkylated with 100 μL of 55 mM iodoacetamide in 100 mM Ambic at room temperature in the dark for 60 min without a traditional reduction step. After wash steps as described above, the gel slices were dried and rehydrated with 50 μL chymotrypsin in 50 mM Ambic, 10% ACN (20 ng/μL) at 37 °C for 16 hrs. The digested peptides were extracted twice with 70 μL of 50% ACN, 5% FA and once with 70 μL of 90% ACN, 5% FA. Extracts from each sample were combined and lyophilized.

The in-gel tryptic digests were reconstituted in 20 μL of 0.5% FA for nanoLC-ESI-MS/MS analysis, which was carried out using an Orbitrap FusionTM TribridTM (Thermo-Fisher Scientific, San Jose, CA) mass spectrometer equipped with a nanospray Flex Ion Source and coupled with a Dionex UltiMate3000RSLCnano system (Thermo, Sunnyvale, CA).[59,60] The gel extracted peptide samples (5–15 μL) were injected onto a PepMap C-18 RP nanotrapping column (5 μm, 100 μm i.d x 20 mm) at 20 μL/min flow rate for rapid sample loading and then separated on a PepMap C-18 RP nano column (2 μm, 75 μm x 25 cm) at 35 °C. The tryptic peptides were eluted in a 90 min gradient of 5% to 38% CAN in 0.1% formic acid at 300 nL/min, followed by a 7 min ramping to 90% ACN-0.1% FA and an 8 min hold at 90% ACN-0.1% FA. The column was re-equilibrated with 0.1% FA for 25 min prior to the next run. The Orbitrap Fusion is operated in positive ion mode with spray voltage set at 1.6 kV and source temperature at 275 °C. External calibration for FT, IT and quadrupole mass analyzers was performed. In data-dependent acquisition (DDA) analysis, the instrument was operated using FT mass analyzer in MS scan to select precursor ions followed by 3 second “Top Speed” data-dependent collision-induced dissociation (CID) ion trap MS/MS scans at 1.6 m/z quadrupole isolation for precursor peptides with multiple charged ions above a threshold ion count of 10,000 and normalized collision energy of 30%. Alternatively, the 3 second “top speed” DDA was conducted under CID/electron transfer dissociation (ETD) toggle method, in which a survey MS scan was followed by data-dependent CID (for 2 charged ions only) and ETD (≥ 3 charged ions) for ion trap MS/MS scans with quadrupole isolation for precursor peptides. The normalized collision energy was 30% for CID, and 150 ms for ETD reaction time and 200 ms for maximal ETD reagent injection time were used for ETD analysis.[59] MS survey scans at a resolving power of 120,000 (fwhm at m/z 200), for the mass range of m/z 375–1575. Dynamic exclusion parameters were set at 40 seconds of exclusion duration with +/− 10 ppm exclusion mass width. The activation time was 10 ms for CID analysis. All data were acquired under Xcalibur 3.0 operation software (Thermo-Fisher Scientific).

The DDA raw files for CID MS/MS only were subjected to database searches using Proteome Discoverer (PD) 1.4 and/or PD 2.2 software (Thermo Fisher Scientific, Bremen, Germany) with the Sequest HT algorithm. While the CID-ETD raw data file was searched using PD 2.1 or PD 2.2 software. The database search was conducted against a customized *E coli* database (host) containing engineered (his-tag) FlgE protein or its various mutants with two-missed trypsin cleavage sites allowed for direct identification of non-crosslinking peptides of targeted FlgE. The peptide precursor tolerance was set to 10 ppm and fragment ion tolerance was set to 0.6 Da. Variable modifications on methionine oxidation, carbamidomethyl modification of cysteine, deamindation of asparagine and glutamine, dehydroalanine formation and BME-adduct were monitored. Only high confidence peptides defined by Sequest HT with a 1% FDR by Percolator were considered for the peptide identification. All MS/MS spectra for identified relevant peptides from initial database searching were manually inspected and validated using Xcalibur 3.0. Confident identification of LA crosslink peptides with alkylated and/or BME adduct forms was achieved by manual confirmation of MS precursor ions and their associated CID MS/MS and ETD MS/MS spectra.

Relative quantitation of identified LA crosslinked peptides and their associated non-crosslinked peptides between samples with different treatment was determined by PD-assistant or manual extraction ion chromatograms (XICs). The XICs of each peptide (and its modified forms) with different charge states for each sample were obtained based on precursor ion m/z with mass tolerance of 5 ppm in Xcalibur software. To quantitatively compare peptide abundances in each sample to one another, an internal peptide (184-LPELPEGANR-193) of Td FlgE, which is not involved in the Lal crosslink reaction was used as a reference peptide for normalization of targeted peptide abundances in each sample as previously reported.[7]

- pH titration of Td $FlgE_{FL}$ WT, C178S, and C178T variants

The crosslinking abilities of $FlgE_{FL}$ mutants C178S and C178T were assessed by varying buffer pH and monitoring HMWC formation via SDS-PAGE. Preliminary data suggested that HMWC formation for the C178S mutant was greatly reduced compared to WT $FlgE_{FL}$. Therefore, crosslinking yields were increased by titrating the pH of the XLB, and compared for WT, C178S, and C178T variants. For each sample, 10 μM $FlgE_{FL}$ in XLB pH 7–11 was incubated for 48 hours at 4°C and analyzed by SDS-PAGE. For XLB buffers with a pH > 9, glycine was used in lieu of Tris.

- DTNB treatment and detection of TNB-HSO_2^{-1} leaving group

To confirm DTNB treatment produced the cysteine thiol-containing TNB leaving group, 200 μM $TdFlgE_{T31\text{-}M434}$ was treated with 4 mM (20x) DTNB in 50 mM Tris pH 7.5 for 48 hours at 4 °C. Stock DTNB was prepared to a concentration of 250 mM in 100% ethanol. For the FlgE only control, the corresponding volume of 100% ethanol was added in lieu of DTNB (1.6 μL). For all samples, the final ethanol concentration was 1.6% (v/v). After 48 hours incubation time, the samples were centrifuged at 13,000 rpm for 10 minutes at 4 °C and the protein removed via centrifugation in a 10 kDa MWCO spin-column. The flow through was then collected and flash frozen in liquid nitrogen and stored at −80 °C until HPLC-MS

analysis. Briefly, 1 μL of sample was injected into an RP-HPLC system (Agilent Zorbax Eclipse XDB-C18 2.1 × 150mm 1.8um-particle size) coupled with a ThermoFisher Scientific Q Exactive Hybrid Quadrupole-Orbitrap Mass Spectrometer. Samples were eluted at 0.5 mL/min with a mobile phase of H_2O:ACN, both containing 0.% (v/v) formic acid. Sample components were resolved using the gradient elution pattern: (time/min, acetonitrile (%)) 00:00–02:00 (2%), 02:00–12:30 (2% to 100%), 12:30–18:30 (100%), 18:30–18:36 (100% to 2%), and 18:36–20:00 (2%). All HPLC-MS data were manipulated and analyzed in Xcalibur operation software (Thermo-Fisher Scientific).

- Reaction of FlgE variants with sulfur oxoform specific probes and metal cations

In order to probe the oxidation state of the cysteine thiol prior to forming DHA, the hyperactive Td $FlgE_{FL}$ K336A variant was reacted with dimedone (DD, >99%, 38490, Sigma), γ-butyrolactone (GBA, 20740, Fluka), and hydrogen peroxide. Stock GBA (13.1M) was diluted to a final concentration of 10 mM, 1.0 mM, and 0.1mM in ddH_2O. Stock DD was prepared in 100% ethanol to a final concentration of 10 mM, 1.0 mM, and 0.1 mM. Hydrogen peroxide (H_2O_2, 30% (v/v)) was diluted to a final concentration of 5%. 2.5%, and 1%. For each sample, 6 μM $FlgE_{FL}$ K336A in 1X XLB pH 8.5 was mixed with DD, GBA, or H_2O_2, and incubated for 48 hours at 4°C. Samples were then quenched with SDS loading dye containing excess BME and analyzed via SDS-PAGE. To ascertain the role of various biologically relevant metal cations in the Lal crosslink formation reaction, the ability of each cation to rescue Lal crosslinking in the crosslinking deficient FlgE truncation $FlgE_{N91\text{-}S423}$ was assessed. For each sample, a 20 mM stock was prepared in 18 MΩ H_2O (magnesium sulfate, manganese sulfate, copper (II) sulfate, cobalt (III) chloride, iron (III) sulfate, calcium chloride and zinc sulfate). To each sample 10 μM $FlgE_{N91\text{-}S423}$ in 1X XLB pH 8.5 was mixed with 1 μL cation stock and incubated at 4 °C for 48 hours. The samples were then quenched and analyzed via SDS-PAGE.

Data Availability:

Coordinates and structure files for WT $FlgE_{D2}$, DHA $FlgE_{D2}$ and Lal crosslinked $FlgE_{D1D2:D2}$ have been deposited to the Protein Data Bank with the following accession codes: 6NDW (WT $FlgE_{D2}$), 6NDT (DHA $FlgE_{D2}$) and 6NDX ($FlgE_{D1D2:D2}$ Lal crosslinked dimer). Raw MS data for all mutants and X-ray diffraction images are available from the corresponding author upon reasonable request. Constructs encoding for full-length and truncated Td FlgE variants are available from the corresponding author upon reasonable request.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgements:

This work was supported by NIH grant R35 122535 (B.R.C.), the CBI Training grant T32 GM008500 (M.J.L. and B.R.C.), NIH R01-DE023431 (N.C., M.M., and C.L.), AI078958 (C.L.), and NIH SIG grant 1S10 OD017992-01 (S.Z.). CHESS is supported by the NSF and NIH/NIGMS (DMR-1332208). MacCHESS is supported by NIH/NIGMS (GM-103485). Remote data collection was performed at the NE-CAT beamlines (GM124165) using an Eiger detector (OD021527) at the Advanced Photon Source (DE-AC02-06CH11357). The authors would like to

thank H. Le for assistance with LC-MS experiments, A. Bilwes-Crane for editing the manuscript, the Cornell Proteomic and MS Facility for providing the mass spectrometry data, and E. Anderson and R. Bahwal for technical assistance with MS sample preparation, data acquisition, and analysis.

References:

1. Kang HJ & Baker EN Intramolecular isopeptide bonds: protein crosslinks built for stress? Trends Biochem. Sci 36, 229–237 (2011). [PubMed: 21055949]
2. Walden M, Crow A, Nelson MD & Banfield MJ Intramolecular isopeptide but not internal thioester bonds confer proteolytic and significant thermal stability to the *S. pyogenes* pilus adhesin Spy0125. Proteins Struct. Funct. Bioinforma 82, 517–527 (2014).
3. Kwon H et al. Autocatalytically generated Thr-Gln ester bond crosslinks stabilize the repetitive Ig-domain shaft of a bacterial cell surface adhesin. PNAS 111, 1367–1372 (2014). [PubMed: 24344302]
4. Baker EN, Squire CJ & Young PG Self-generated covalent crosslinks in the cell-surface adhesins of Gram-positive bacteria. Biochem. Soc. Trans 43, 787–794 (2015). [PubMed: 26517883]
5. Popa MP, McKelvey TA, Hempel J & Hendrix RW Bacteriophage HK97 structure: wholesale covalent crosslinking between the major head shell subunits. J. Virol 65, 3227–37 (1991). [PubMed: 1709700]
6. Williams M & Baxter R The structure and function of thioester-containing proteins in arthropods. Biophys. Rev 6, 261–272 (2014). [PubMed: 28510031]
7. Miller MR et al. Spirochaete flagella hook proteins self-catalyse a lysinoalanine covalent crosslink for motility. Nat. Microbiol 1, 16134 (2016). [PubMed: 27670115]
8. Charon NW et al. The unique paradigm of spirochete motility and chemotaxis. Annu. Rev. Microbiol 66, 349–70 (2012). [PubMed: 22994496]
9. Wolgemuth CW Flagellar motility of the pathogenic spirochetes. Semin. Cell Dev. Biol 46, 104–112 (2015). [PubMed: 26481969]
10. Shibata S et al. FliK regulates flagellar hook length as an internal ruler. Mol. Microbiol 64, 1404–1415 (2007). [PubMed: 17542929]
11. Zhao X et al. Cryoelectron tomography reveals the sequential assembly of bacterial flagella in Borrelia burgdorferi. Proc. Natl. Acad. Sci 110, 14390–14395 (2013). [PubMed: 23940315]
12. Kojima S & Blair DF in International review of cytology 233, 93–134 (2004). [PubMed: 15037363]
13. Samatey FA et al. Structure of the bacterial flagellar hook and implication for the molecular universal joint mechanism. Nature 431, 1062–1068 (2004). [PubMed: 15510139]
14. Zhao X, Norris SJ & Liu J Molecular architecture of the bacterial flagellar motor in cells. Biochemistry 53, 4323–33 (2014). [PubMed: 24697492]
15. Alegre-Cebollada J, Badilla CL & Fernández JM Isopeptide Bonds Block the Mechanical Extension of Pili in Pathogenic *Streptococcus pyogenes*. J. Biol. Chem 285, 11235–11242 (2010). [PubMed: 20139067]
16. Dierkes LE, Peebles CL, Firek BA, Hendrix RW & Duda RL Mutational analysis of a conserved glutamic acid required for self-catalyzed crosslinking of bacteriophage HK97 capsids. J. Virol 83, 2088–2098 (2009). [PubMed: 19091865]
17. Ross PD et al. Crosslinking renders bacteriophage HK97 capsid maturation irreversible and effects an essential stabilization. EMBO J 24, 1352–1363 (2005). [PubMed: 15775971]
18. Duda RL et al. Structure and energetics of encapsidated DNA in bacteriophage HK97 studied by scanning calorimetry and cryo-electron microscopy. J. Mol. Biol 391, 471–83 (2009). [PubMed: 19540242]
19. Friedman M Lysinoalanine in food and in antimicrobial proteins. Adv. Exp. Med. Biol 459, 145–59 (1999). [PubMed: 10335374]
20. Repka LM, Chekan JR, Nair SK & van der Donk WA Mechanistic Understanding of Lanthipeptide Biosynthetic Enzymes. Chem. Rev 117, 5457–5520 (2017). [PubMed: 28135077]
21. Huo L, Ökesli AA, Zhao M & Van Der Donk WA Insights into the Biosynthesis of Duramycin. Appl. Environ. Microbiol 83, 2698–2714 (2017).

22. Ökesli A, Cooper LE, Fogle EJ & van der Donk WA Nine Post-translational Modifications during the Biosynthesis of Cinnamycin. J. Am. Chem. Soc 133, 13753–13760 (2011). [PubMed: 21770392]
23. An L et al. Substrate-assisted enzymatic formation of lysinoalanine in duramycin. Nat. Chem. Biol 14, 928–933 (2018). [PubMed: 30177849]
24. Miller KA et al. Initial characterization of the FlgE hook high molecular weight complex of Borrelia burgdorferi. PLoS One 9, e98338 (2014). [PubMed: 24859001]
25. Lukszo J, Patterson D, Albericio F & Kates SA 3-(1-Piperidinyl)alanine formation during the preparation of C-terminal cysteine peptides with the Fmoc/t-Bu strategy. Lett. Pept. Sci 3, 157–166 (1996).
26. Imada K Bacterial flagellar axial structure and its construction. Biophys. Rev 10, 559–570 (2018). [PubMed: 29235079]
27. Homma M, DeRosier DJ & Macnab RM Flagellar hook and hook-associated proteins of Salmonella typhimurium and their relationship to other axial components of the flagellum. J. Mol. Biol 213, 819–832 (1990). [PubMed: 2193164]
28. Moriya N, Minamino T, Hughes KT, Macnab RM & Namba K The Type III Flagellar Export Specificity Switch is Dependent on FliK Ruler and a Molecular Clock. J. Mol. Biol 359, 466–477 (2006). [PubMed: 16630628]
29. Moriya N et al. Role of the Dc domain of the bacterial hook protein FlgE in hook assembly and function. Biophysics (Oxf) 9, 63–72 (2013).
30. Fujii T, Kato T & Namba K Specific arrangement of alpha-helical coiled coils in the core domain of the bacterial flagellar hook for the universal joint function. Structure 17, 1485–93 (2009). [PubMed: 19913483]
31. Matsunami H, Barker CS, Yoon Y-H, Wolf M & Samatey FA Complete structure of the bacterial flagellar hook reveals extensive set of stabilizing interactions. Nat. Commun 7, 13425 (2016). [PubMed: 27811912]
32. Yoon Y-H et al. Structural insights into bacterial flagellar hooks similarities and specificities. Sci. Rep 6, 35552 (2016). [PubMed: 27759043]
33. Grabarek Z & Gergely J Zero-length crosslinking procedure with the use of active esters. Anal. Biochem 185, 131–135 (1990). [PubMed: 2344038]
34. Brennan DF & Barford D Eliminylation: a post-translational modification catalyzed by phosphothreonine lyases. Trends Biochem. Sci 34, 108–114 (2009). [PubMed: 19233656]
35. Zhou G, Wang H, Ma Y & Chen X An NBD fluorophore-based colorimetric and fluorescent chemosensor for hydrogen sulfide and its application for bioimaging. Tetrahedron 69, 867–870 (2013).
36. Lynch MJ & Crane BR Design, validation, and application of an enzyme coupled hydrogen sulfide detection assay. Biochemistry 59, 474–483 (2019).
37. Samatey FA et al. Structure of the bacterial flagellar hook and implication for the molecular universal joint mechanism. Nature 431, 1062–1068 (2004). [PubMed: 15510139]
38. Kido Y, Yoon YH & Samatey FA Crystallization of a 79 kDa fragment of the hook protein FlgE from Campylobacter jejuni. Acta Crystallogr. Sect. F. Struct. Biol. Cryst. Commun 67, 1653–7 (2011).
39. Chalker JM et al. Methods for converting cysteine to dehydroalanine on peptides and proteins. Chem. Sci 2, 1666 (2011).
40. Galan SRG et al. Post-translational site-selective protein backbone α-deuteration. Nat. Chem. Biol 14, 955–963 (2018). [PubMed: 30224694]
41. Glavas S & Tanner ME Catalytic acid/base residues of glutamate racemase. Biochemistry 38, 4106–4113 (1999). [PubMed: 10194325]
42. Shaikh TR et al. A partial atomic structure for the flagellar hook of Salmonella typhimurium. Proc. Natl. Acad. Sci 102, 1023–1028 (2005). [PubMed: 15657146]
43. Bharadwaj KC Acryl Activation by Intramolecular Hydrogen Bond: Morita Baylis Hillman Reaction of Acrylamide with Broad Substrate Scope. Chem. Sel 2, 5384–5389 (2017).

44. Chevance FFV & Hughes KT Coordinating assembly of a bacterial macromolecular machine. Nat. Rev. Microbiol 6, 455–465 (2008). [PubMed: 18483484]
45. Zhao X et al. Cryoelectron tomography reveals the sequential assembly of bacterial flagella in Borrelia burgdorferi. Proc. Natl. Acad. Sci 110, 14390–14395 (2013). [PubMed: 23940315]
46. Roberts MC, Chung WO & Roe DE Characterization of tetracycline and erythromycin resistance determinants in Treponema denticola. Antimicrob. Agents Chemother 40, 1690–4 (1996). [PubMed: 8807063]
47. Mitjà O et al. Re-emergence of yaws after single mass azithromycin treatment followed by targeted treatment: a longitudinal study. Lancet 391, 1599–1607 (2018). [PubMed: 29428183]
48. Nikolaidis I, Favini-Stabile S & Dessen A Resistance to antibiotics targeted to the bacterial cell wall. Protein Sci 23, 243–259 (2014). [PubMed: 24375653]
49. Bernardes GJL, Chalker JM, Errey JC & Davis BG Facile Conversion of Cysteine and Alkyl Cysteines to Dehydroalanine on Protein Surfaces: Versatile and Switchable Access to Functionalized Proteins. J. Am. Chem. Soc 130, 5052–5053 (2008). [PubMed: 18357986]
50. Dadová J, Galan SR & Davis BG Synthesis of modified proteins via functionalization of dehydroalanine. Curr. Opin. Chem. Biol 46, 71–81 (2018). [PubMed: 29913421]
51. Schindelin J et al. Fiji: an open-source platform for biological-image analysis. Nat. Methods 9, 676–682 (2012). [PubMed: 22743772]
52. Adams PD et al. PHENIX: a comprehensive Python-based system for macromolecular structure solution. Acta Crystallogr. Sect. D Biol. Crystallogr 66, 213–221 (2010). [PubMed: 20124702]
53. Samatey FA et al. Structure of the bacterial flagellar hook and implication for the molecular universal joint mechanism. Nature 431, 1062–8 (2004). [PubMed: 15510139]
54. Emsley P, Lohkamp B, Scott WG & Cowtan K Features and development of Coot. Acta Cryst 66, 486–501 (2010).
55. Emsley P et al. Coot: model-building tools for molecular graphics. Acta Crystallogr. Sect. D Biol. Crystallogr 60, 2126–2132 (2004). [PubMed: 15572765]
56. The PyMOL Molecular Graphics System, Version 2.0 Schrödinger, LLC.
57. Otwinowski Z & Minor W Processing of X-ray diffraction data collected in oscillation mode. Methods Enzymol 276, 307–326 (1997).
58. Yang Y, Thannhauser TW, Li L & Zhang S Development of an integrated approach for evaluation of 2-D gel image analysis: Impact of multiple proteins in single spots on comparative proteomics in conventional 2-D gel/MALDI workflow. Electrophoresis 28, 2080–2094 (2007). [PubMed: 17486657]
59. Thomas CJ, Cleland TP, Zhang S, Gundberg CM & Vashishth D Identification and characterization of glycation adducts on osteocalcin. Anal. Biochem 525, 46–53 (2017). [PubMed: 28237256]
60. Yang Y, Anderson E & Zhang S Evaluation of six sample preparation procedures for qualitative and quantitative proteomics analysis of milk fat globule membrane. Electrophoresis 39, 2332–2339 (2018). [PubMed: 29644703]

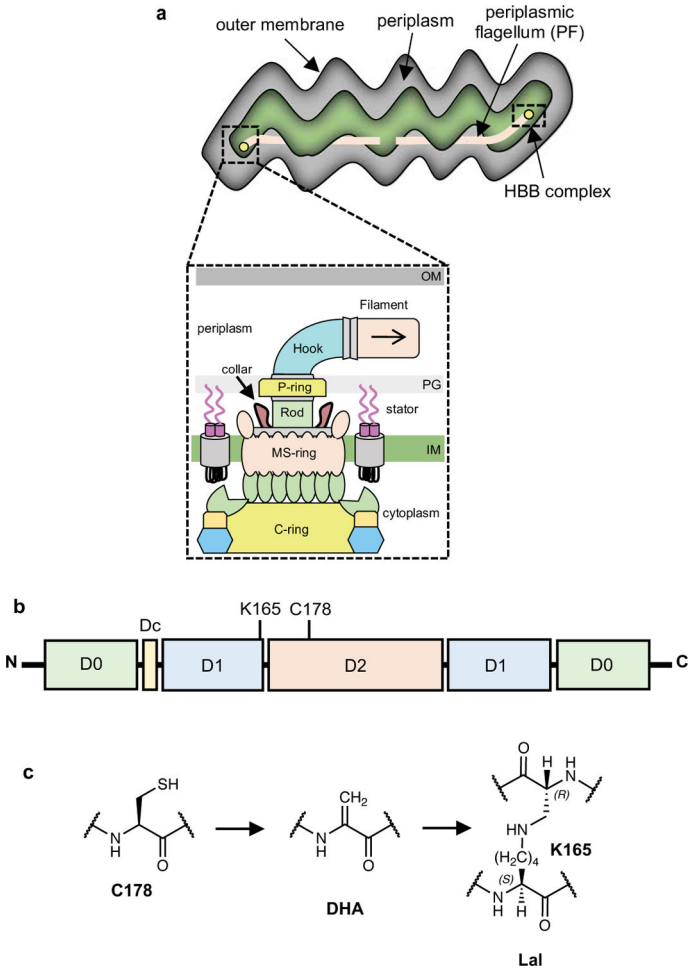


Fig. 1: Overview of lysinoalanine crosslinking reaction catalyzed by Td FlgE.

a, Cartoon of a spirochaete cell highlighting the location of the outer membrane, periplasm, periplasmic flagellum (PF), and hook-basal body (HBB) complex. The inset emphasizes the location of the hook (blue) relative to the outer (OM)/inner (IM) membranes, peptidoglycan layer (PG), and other miscellaneous flagellar motor components (filament, P-ring, rod, L-ring, MS-ring, C-ring, stators, and collar). **b,** Domain organization of Td FlgE with the C178 and K165 positions marked. **c,** Lal crosslink formation occurs via the conversion of C178 to DHA. Aza-Michael addition of the side chain ε-amine of K165 to the DHA Cβ yields the mature Lal crosslinkage.

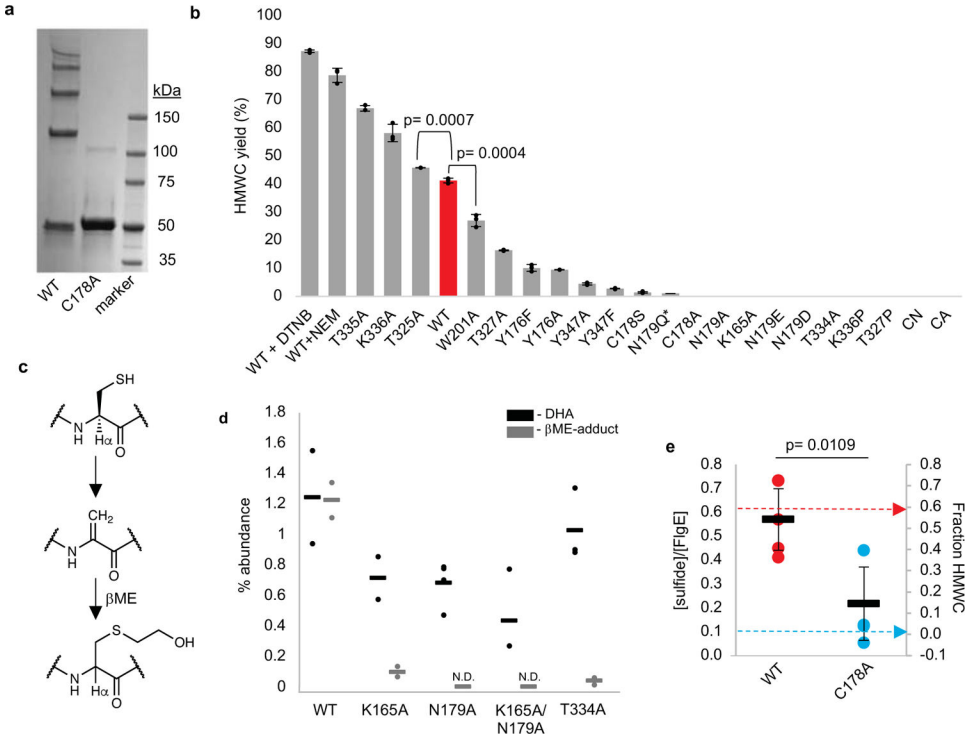


Fig. 2: Residue requirements for Lal crosslinking.

a, Representative SDS-PAGE gel showing Lal crosslinked HMWC formation in WT but not C178A Td $FlgE_{FL}$. Experiments were repeated at least three times with similar results. **b,** Quantification of percent HMWC formation by $FlgE_{FL}$ variants. Histogram bars represent the mean value of three biological replicates. Error bars are reported as ± the standard deviation of n = 3 independent replicates for each variant. Raw SDS-PAGE gels used to determine percent HMWC for all Td FlgE variants are reported in Supplementary Figure 11b.* The N179Q variant concentration was increased to 20 μM, and the percent HMWC formation determined from a single biological replicate. **c,** Chemical structure of the modified FlgE peptides. **d,** Relative amounts of DHA- and βME adduct-containing peptides in WT, N179A, K165A, N179A/K165A, and T334 $FlgE_{FL}$. For each variant, bars represent the average percent DHA-containing peptide (black) and βME-adduct-containing peptide (gray). Technical replicates for each variant (WT, n = 2; K165A, n = 2; N179A, n = 4; K165A/N179A, n = 3; T334A, n = 3) are represented as individual data points (N.D. = not detected). Representative extracted ion chromatograms (XICs) of each peptide for all FlgE variants are reported in Supplementary Figure 12. **e,** Quantification of sulfide release by Td FlgE Lal crosslinking using the NBD-N_3 sulfide detection assay.[17] For each sample (red, WT; blue, C178A), the circular data points represent the results of each individual trial, and the mean value of four biological replicates denoted with a black bar. Error bars are reported as ± the standard deviation of n = 4 independent trials. Red and blue dotted-arrows signify the total HMWC yield for WT and C178 Td $FlgE_{FL}$, respectively, measured in **a**. Statistical significance determined by an unpaired two-tailed Student's t-test for results reported in **b** and **e**. See Supplementary Figure 13 for an uncropped image of the SDS-PAGE gel reported in **a**.

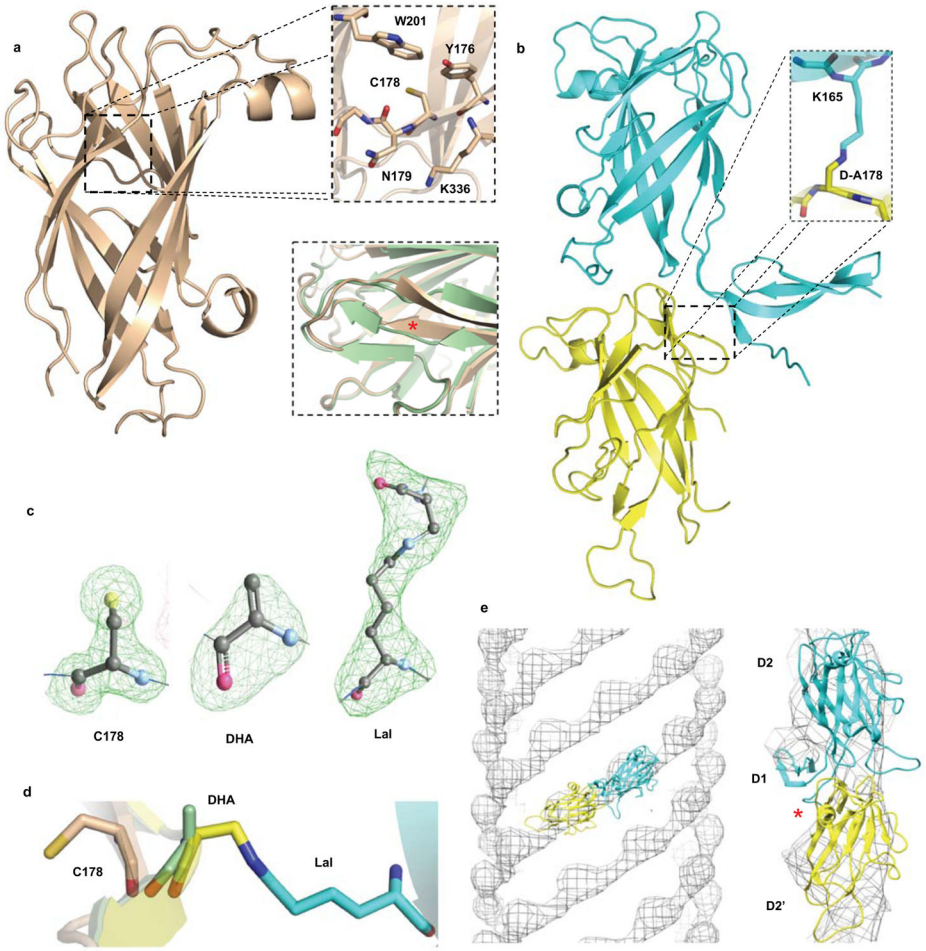


Fig. 3: Comparison of pre-crosslink, dehydroalanine and post-crosslink FlgE.

a, Structure of Td $FlgE_{D2}$ (tan). Pre-crosslinking C178 is caged within the β-barrel by Y176, N179, W201, and K336 (top inset). Conversion of C178 to DHA (green ribbons) causes a localized melting of secondary structure (bottom inset, RMSD: 0.472, 131 residues, location of DHA indicated with red asterisk). **b,** Structure of the Lal crosslinked $FlgE_{D1D2:D2}$ dimer ($FlgE_{D157\text{-}I371}$:$FlgE_{A186\text{-}E339}$). The Lal crosslink (inset) lies at the interface between a $FlgE_{D2}$ domain (yellow) and the linker to the $FlgE_{D1D2}$ domain (cyan). **c,** Unbiased omit maps for C178, DHA, and Lal crosslink (mFo-DFc, map contour = 1.5 σ). **d,** Superimposition of the backbone and sidechain re-orientation associated with C178, DHA, and Lal crosslink structures demonstrates the structural basis for the observed inversion of stereochemistry about the 178 Cα position. **e.** Superimposition of the Lal crosslinked FlgE dimer into intact hook density from *Salmonella*[20] (approximate location of Lal crosslink is indicated with red asterisk).

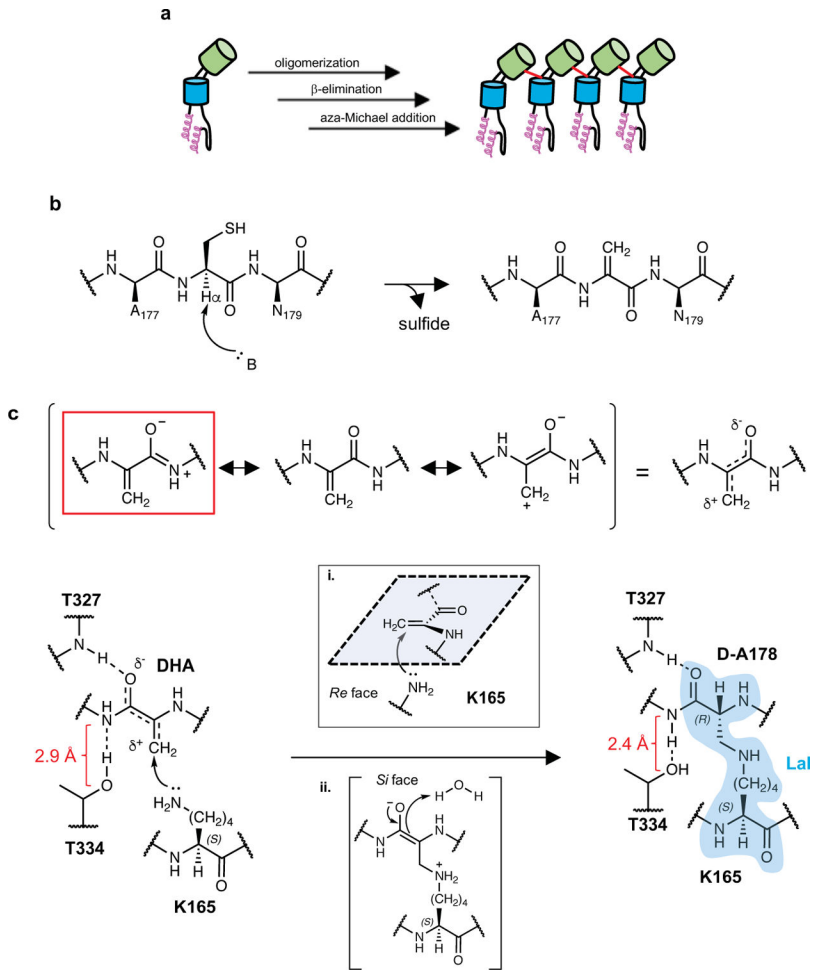


Fig. 4: Proposed model of Lal crosslink formation in Td FlgE.

a, Lal crosslink formation proceeds by three distinct biochemical steps: oligomerization, β-elimination, and aza-Michael addition. Pink helices, D0 domains; blue cylinders, D1 domains; green cylinders, D2 domains; red lines, Lal crosslinks. **b,** Cysteine to DHA conversion is triggered by the general base (denoted by B:)-catalyzed deprotonation of the C178 Hα. DHA formation occurs via the β-elimination of C178 and releases hydrogen sulfide (H_2S/HS^-) as a byproduct. **c,** Overview of the aza-Michael addition of K165 from an adjacent FlgE monomer to yield Lal. Top, hydrogen bonding with T334 polarizes the DHA peptide backbone to disfavor the imine-containing DHA resonance form (indicated by a red box). Bottom, attack of the ε-amine of K165 from the *Re* face (inset *i*) yields an enolate that is stabilized by a hydrogen bond with the backbone amide of T327. Keto-enol tautomerization of the enolate coupled with protonation from solvent on the *Si* face (inset *ii*) affords the mature Lal crosslink (blue) with the observed *(R)* or D-alanine stereochemistry. The solvent molecules depicted in **c** (inset *ii*) are distinct from those observed in the DHA and Lal crosslinked structures reported in Supplementary Figure 8. DHA to Lal conversion is coupled to a decrease in the hydrogen bond distance from 2.9 Å to 2.4 Å between T334 and the backbone amide nitrogen of N179, suggesting a switch of T334 from hydrogen bond donor to acceptor and an increase in the imine character of the Lal178-N179 peptide bond.