SCIENTIFIC REP🅞RTS

Engineering a Lys-Asn isopeptide bond into an immunoglobulin-like protein domain enhances its stability

Hanna Kwon*, Paul G. Young*, Christopher J. Squire & Edward N. Baker

The overall stability of globular protein structures is marginal, a balance between large numbers of stabilizing non-covalent interactions and a destabilizing entropic term. Higher stability can be engineered by introduction of disulfide bonds, provided the redox environment is controlled. The discovery of stabilizing isopeptide bond crosslinks, formed spontaneously between lysine and asparagine (or aspartic acid) side chains in certain bacterial cell-surface proteins suggests that such bonds could be introduced by protein engineering as an alternative protein stabilization strategy. We report the first example of an isopeptide bond engineered *de novo* into an immunoglobulin-like protein, the minor pilin FctB from *Streptococcus pyogenes*. Four mutations were sufficient; lysine, asparagine and glutamic acid residues were introduced for the bond-forming reaction, with a fourth Val/Phe mutation to help steer the lysine side chain into position. The spontaneously-formed isopeptide bond was confirmed by mass spectrometry and X-ray crystallography, and was shown to increase the thermal stability by 10 °C compared with the wild type protein. This novel method for increasing the stability of IgG-like proteins has potential to be adopted by the field of antibody engineering, which share similar β-clasp Ig-type domains.

Proteins have evolved, in general, to have only marginal stability, with their folded structures representing a balance between large stabilizing and destabilizing free energy terms. An intriguing aspect of their evolution is that despite the chemical diversity inherent in the side chains of their constituent amino acids, chemical reactions between these side chains are very rare. Evidently their sequences have evolved in such a way as to avoid reactions that might prejudice normal folding. Thus the only common covalent linkages between protein side chains are disulfide bonds, and inappropriate formation of these bonds is kept in check by the redox environment and by specialized enzyme machinery.

A striking exception to this paradigm has been the discovery, in many cell-surface proteins of Gram-positive bacteria, of covalent bond crosslinks that result from autocatalytic reactions between amino acid side chains during protein folding[1–3]. Intramolecular isopeptide bonds, formed between the ε-amino group of a Lys side chain and the carboxamide group of an Asn side chain[1] or, less often, the carboxylate group of an Asp side chain[4], are widely present in the pilin subunits of Gram-positive bacterial pili[1,2,5–16] and in the repetitive stalk domains of multidomain adhesins[4,17,18]. Autocatalytic formation of these bonds, in a single-turnover event, depends on an adjacent acidic residue, Glu or Asp, that acts as a proton shuttle[1,4,19]. The evident role of the crosslinks is to confer increased mechanical, thermal and proteolytic stability[20,21], assisting these long, thin protein structures to withstand the chemical and mechanical stresses encountered when engaging host receptors[22]. A second type of crosslink, involving an ester bond between Thr and Gln side chains, has also been discovered and characterized recently in the stalk domains of bacterial cell-surface adhesins[3]. Sequence searches suggest that bonds of this kind are also widespread.

The stability and strength of isopeptide bonds, and the ability to form them autocatalytically in a tightly controlled environment makes them attractive for applications in biotechnology. A variety of creative applications have been developed from natural isopeptide-containing domains derived from the pilus subunit Spy0128[23] and

Maurice Wilkins Centre for Molecular Biodiscovery and School of Biological Sciences, University of Auckland, Private Bag 92019, Auckland, New Zealand. *These authors contributed equally to this work. Correspondence and requests for materials should be addressed to C.J.S. (email: c.squire@auckland.ac.nz) or E.N.B. (email: ted.baker@auckland.ac.nz)
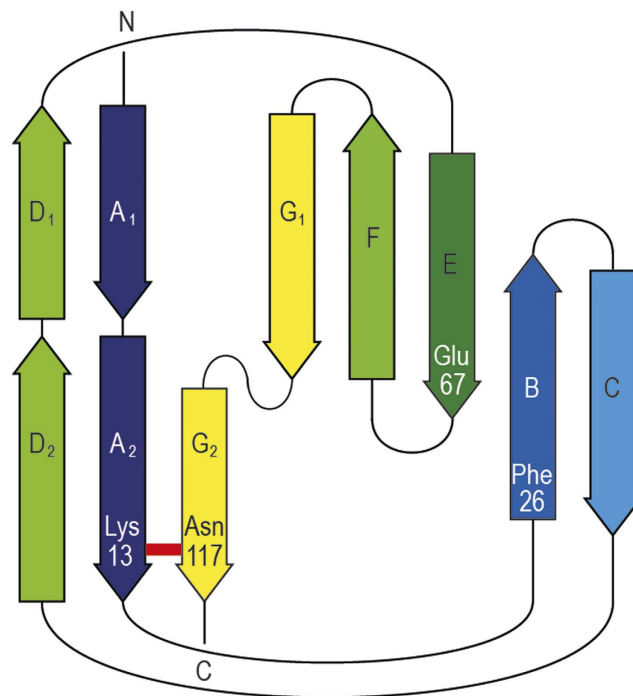
Figure 1. Topology diagram for FctB-iso showing the sites of the residues mutated. The red bar indicates an isopeptide bond between Lys13 and Asn117, linking the first and last strands of the IgG-like fold.

the adhesin FbaB[24], both from *Streptococcus pyogenes*. Spontaneous isopeptide bond formation has been demonstrated to occur when these domains are split into a peptide carrying the Asn/Asp residue and a receptor with the complementary Lys and Glu residues.

The potential also exists to generate proteins with greatly enhanced thermostability by engineering them with the ability to form isopeptide bonds on folding, analogous to the way in which proteins have been stabilized by introduction of non-native disulfide bonds[25,26]. Engineering isopeptide bonds is potentially more challenging, as it requires the involvement of three residues rather than two, and the steric and environmental requirements for the bond-forming reaction are not yet well understood. The advantages of crosslinks of this kind are considerable, however, since they would not be affected by the redox environment, as disulfide bonds are.

Here we describe the first successful introduction of an isopeptide bond into a non-isopeptide containing protein. The host protein is FctB, the basal subunit in *S. pyogenes* pili[27]. This single-domain protein has the same CnaB-type fold as is found for many isopeptide-containing domains, but has none of the required Lys, Glu and Asn residues at the usual crosslinking site. It also has substantial structural differences, with only 15% sequence identity and a root-mean-square difference (rmsd) of 1.8 Å over 96 Cα positions when compared with its closest homolog, the N-terminal domain of Spy0128[27]. We show that by introduction of Lys, Glu and Asn residues into the appropriate positions in the protein fold, together with one further mutation to generate steric constraints on the Lys residue, spontaneous isopeptide bond formation can be achieved. The stability of the domain is substantially enhanced, raising the possibility of engineering crosslinks of this kind into other Ig-like domains as the "rules" governing bond formation become better understood.

Results

Isopeptide design and engineering. Structural alignment of FctB[27] with its closest isopeptide-containing homolog, the N-terminal domain of Spy0128 (Spy0128-N), guided the design of mutations aimed at engineering an isopeptide bond into FctB. The isopeptide bond in Spy0128-N is formed between Lys-36 on its first β-strand (βA) and Asn-168 on its last β-strand (βG), with the spontaneous intramolecular reaction being catalyzed by Glu-117 on the fifth β-strand (βE)[1]. In the archetypal CnaB fold shared by FctB and Spy0128, βA and βG are adjacent parallel strands in the 3-stranded β–sheet whereas βE is a central strand in the opposing 4-stranded β-sheet (Fig. 1). With low sequence identity (15%) between FctB and Spy0128-N, the two β-sheets do not pack together identically, a phenomenon common in Ig-superfamily domains. Nevertheless, a structural superposition identified Asn-13 and Pro-117 in FctB as structurally homologous with the isopeptide bond forming residues Lys-36 and Asn-168 in Spy0128-N and Gln-67 in FctB was identified as the analog of the catalytic Glu-117[27].

We first produced a construct of FctB containing the three minimal mutations (Asn13Lys, Gln67Glu and Pro117Asn) by three successive rounds of site-directed mutagenesis. The triple mutant protein was expressed and purified, and analysed by mass spectrometry, but was found not to have formed an isopeptide bond. We then turned to a closer examination of the environment of the three putative bond-forming residues. The residues surrounding the isopeptide bond in Spy0128-N are almost all hydrophobic, and most of those in closest contact are matched by similar residues in FctB; Leu-38 by Leu-15, Phe-52 by Phe-24, Val-81 by Ile-42 and Phe-166 by
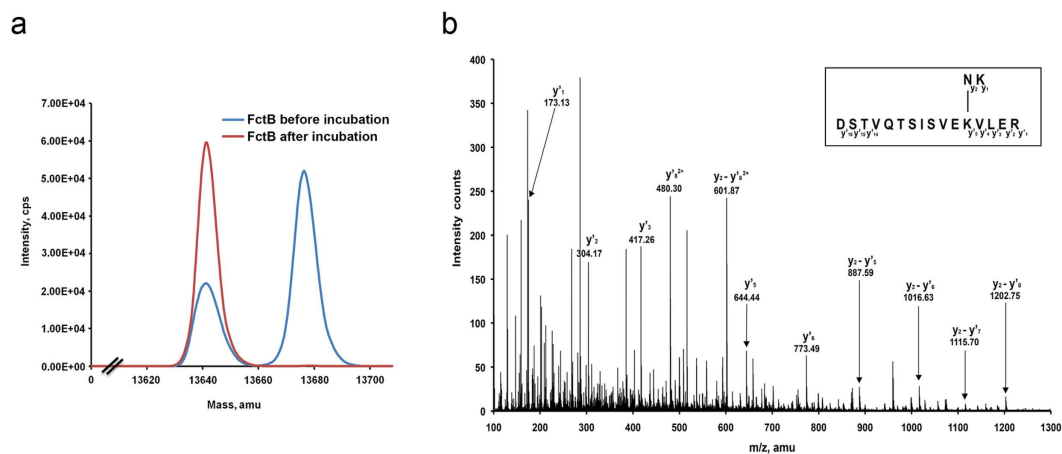
Figure 2. Characterization of isopeptide bond formation by mass spectrometry. (a) Overlay of traces showing the molecular mass measured by ESI-TOF mass spectrometry before incubation (blue) and after incubation for 48 h at 20 °C (red). The blue line shows the two forms of the FctB-iso protein immediately after its expression and purification (Mr 13,648 Da and 13,665 Da). The red line shows that the population shifts to a single isopeptide bond formed species (Mr 13,648 Da) after incubation. **(b)** MS/MS spectra of peptides containing the intramolecular isopeptide bond in FctB-iso. The MS/MS fragment peaks are labelled with *m/z* and the corresponding residues are shown in the inset.

Phe-115. One significant difference was apparent, however. In Spy0128-N, the aliphatic portion of the Lys-36 side chain is sandwiched between Phe-166 on one side and Phe-54 on the other, making van der Waals contacts (4.0–4.4 Å) with both. In FctB, however, Phe-54 is replaced by the much smaller Val-26. We hypothesized that a Val-26 to Phe mutation would more closely mimic the Spy0128 structure, and would restrict the movement of the engineered Lys-13, enclosing it in a more hydrophobic environment and directing it into closer contact with Asn-117.

The quadruple mutant N13K/V26F/Q67E/P117N (hereinafter referred to as FctB-iso) was made, and subjected to analysis by mass spectrometry. This showed that immediately after protein purification a mixed population, with and without isopeptide bond, was present but that bond formation became complete after incubation for 48 hours at 20 °C (Fig. 2a). The molecular weight of FctB-iso was measured as 13.648 kDa, 17 Da less than the calculated mass of FctB (13.665 kDa) and consistent with a loss of NH_3 accompanying formation of a Lys-Asn isopeptide bond (Fig. 2a). Confirmation of the identity of the bond-forming residues was provided by trypsin digestion coupled with liquid chromatography-tandem mass spectrometry (LC-MS/MS)[28]. Manual inspection of the observed fragmentation pattern enabled a peptide fragment to be identified with a covalent isopeptide linkage between Lys-13 and Asn-117 (Fig. 2b).

Protein structure. Crystallization trials on the FctB-iso construct were unsuccessful. We turned, therefore, to a fusion protein in which FctB-iso was joined via a flexible 4-residue linker (Ala-Gly-Gly-Ala) to the C-terminus of maltose binding protein (MBP). Crystals of this His$_6$-MBP-FctB-iso construct were obtained that diffracted to 2.0 Å resolution. The crystal structure was solved by molecular replacement using the MBP[29] and WT FctB[27] structures as models and refined to an R factor of 19.2% and R_{free} of 23.5% (Table 1). The asymmetric unit contains two molecules, A and B. For Molecule A, both the MBP and FctB-iso moieties could be modelled, but Molecule B lacked interpretable electron density for FctB-iso. We presume this is due to disorder; there was no evidence of twinning, incorrect space group, or any anomaly in the data. Examination of the crystal packing shows that the FctB-iso domain in molecule B would be positioned adjacent to a large solvent channel where it has freedom to move on its flexible MBP linker.

The crystal structure contained well-defined electron density that extended over the side chains of Lys-13 and Asn-117 (Fig. 3a), clear confirmation of the presence of the covalent linkage between these residues. The introduction of the isopeptide bond does not appear to cause any perturbation of the overall polypeptide conformation, as superposition of the FctB-iso structure on to that of the parent FctB gives an rmsd of 0.73 Å over 118 Cα atoms. As in Spy0128, the isopeptide has a *cis*-configuration, with its C=O and NH groups hydrogen bonded to the side chain carboxyl group of the adjacent (engineered) Glu-67 residue (Fig. 3a). The isopeptide plane is stacked against the aromatic ring of Phe-24, ~3.4 Å away (Fig. 3b), in almost identical fashion to the isopeptide in Spy0128-N, which stacks with Phe-52. The aliphatic part of the introduced Lys-13 side chain is in van der Waals contact with surrounding hydrophobic side chains; Cγ is 4.2 Å from Ile42, Cδ contacts Phe-115 (4.3 Å) and the introduced Phe-26 (3.9 Å) and Cε contacts Phe-24 (3.4 Å) and Leu-15 (3.5 Å).

One difference in FctB-iso is that its catalytic Glu residue, Glu-67, has two discrete conformations, with its carboxyl group rotated ~90° in one conformation relative to the other. In contrast, the equivalent residue in Spy0128-N has just a single, well-defined conformation. The disorder does not affect its interaction with the isopeptide, however, as both conformers allow the Glu-67 carboxyl oxygens to hydrogen bond with the isopeptide C=O and NH groups (Fig. 3). These interactions appear almost equally favourable in each case, and the only

	FctB-iso
Data collection	Australian synchrotron
X-ray source	MX2
Wavelength	0.9537
Resolution[a] (Å)	19.57–2.00 (2.11–2.00)
Space group	C 222$_1$
Cell dimensions (Å)	a = 180.36, b = 192.09, c = 68.20
Total reflections[a]	584724 (85160)
Unique reflections[a]	78932 (11319)
Redundancy[a]	7.4 (7.4)
Completeness[a] (%)	98.5 (97.8)
R_{pim}[a]	0.053 (0.577)
$I/\sigma I$[a]	10.3 (1.6)
CC1/2[a]	0.997 (0.525)
Refinement	
Resolution (Å)	19.57–2.00
R_{work}/R_{free}	0.194/0.231
No. atoms	
Protein	6547
Maltotriose	68
Water	280
Average B-factors (Å^2)	
Protein	35.9
Maltotriose	36.3
Water	39.7
Rms deviations	
Bond lengths (Å)	0.007
Bond angles (°)	1.14
Ramachandran favored (%)	96.5
Ramachandran outliers (%)	0
PDB code	5ttd

Table 1. Data collection and refinement statistics. [a]Numbers in parentheses are for the outermost shell. [b]CC1/2 = Correlation coefficient[43]. [c]R_{pim} = Precision-indicating R factor[44].

differences are that in conformer 1 Oε2 is in non-bonded contact (3.0 Å) with Ala-77 Cβ, whereas in conformer 2 it has an additional hydrogen bond with the peptide NH of Asp-78.

Thermal stability. The contribution of the engineered isopeptide bond to the thermal stability of FctB-iso was investigated by both differential scanning fluorimetry (DSF) and circular dichroism (CD) spectroscopy. In the DSF method, the global unfolding of a protein is measured by heating the sample from 25 °C to 95 °C and monitoring the exposure of hydrophobic surfaces with a dye that emits strong fluorescence at 600 nm when placed in a non-polar environment[30]. Both wild type and FctB-iso proteins showed a single unfolding transition (melting temperature, T_m) consistent with properly folded proteins. T_m for the engineered FctB-iso was 79 °C, an increase of 9 °C compared with the wild type protein (Fig. 4a). A similar analysis using CD spectroscopy, which enables denaturation to be observed directly, gave a very similar result, revealing a 10 °C difference between WT ($T_m = 71$ °C) and FctB-iso ($T_m = 81$ °C) samples (Fig. 4b).

Discussion

The isopeptide bond forming reaction has been utilised in several engineering applications, but these make use of "deconstructed" isopeptide-containing domains in which the native sequence (and presumably the tertiary structure) around the bond-forming site are kept identical[23,24]. The domain is split into two parts, a receptor domain lacking its final β-strand that carries the key Asn/Asp residue, and a peptide that contains the sequence of the missing strand. Experience with these systems suggested that the local environment was very sensitive to change, as shown by the observation that any changes to the sequence of the final β-strand were deleterious to the bond-forming reaction[24].

Our results indicate that two key factors are essential for isopeptide bond formation: hydrophobicity and proximity. Theoretical studies have shown that a hydrophobic environment is necessary in order to lower the pKa of the Lys ε-amino group, favouring its non-protonated form and enabling it to mount a nucleophilic attack on Cγ of the recipient Asn/Asp acid side chain[4,19]. The hydrophobic environment also raises the pKa of the associated Glu residue, enabling it to act as a proton shuttle in the autocatalytic bond-forming reaction.

We have shown here, however, that proximity of the reacting residues is equally important. The flexible lysine side chain, with its many degrees of freedom, must be steered into position by strategically placed hydrophobic

a

b
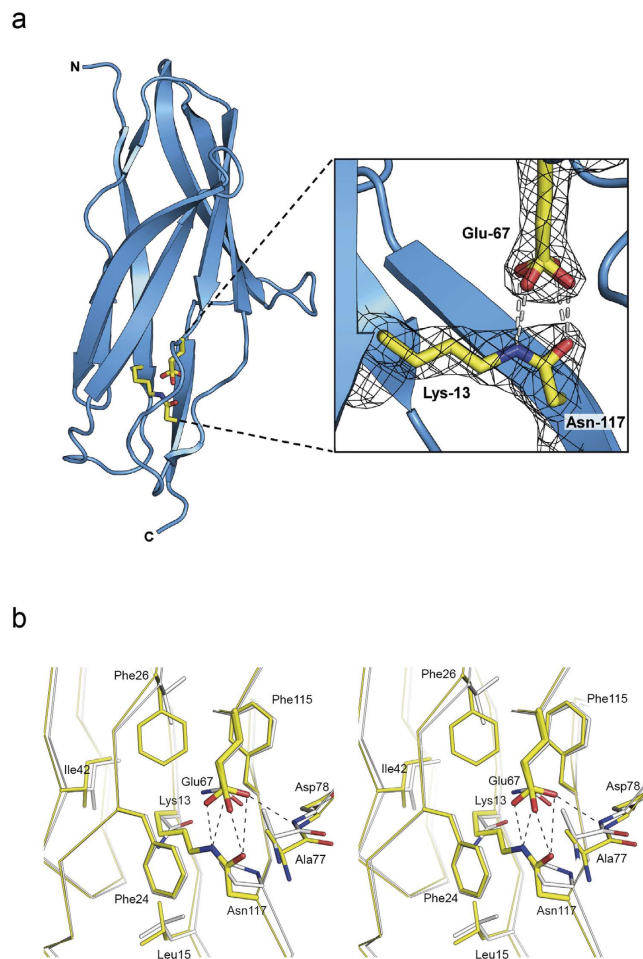
Figure 3. Structure of FctB-iso. (**a**) Ribbon diagram of FctB-iso showing the position of the engineered isopeptide bond. Residues involved in isopeptide bond formation are shown in stick form (yellow; Lys-13, Glu-67 and Asn-117) and are encompassed in bias-removed $2F_o − F_c$ density contoured at 1.5 σ (inset). The carboxyl group of Glu-67 has two discrete conformations, differing by a ~90° rotation about Cγ-Cδ, but with each conformer enabling two hydrogen bonds (dashed lines) between the carboxyl oxygens and the NH and C=O groups of the *cis*-isopeptide moiety. (**b**) Stereo superimposition of the engineered residues in FctB-iso (yellow; Lys-13, Phe-26, Glu-67 and Asn-117) with corresponding residues in WT FctB (white; Asn-13, Val-26, Gln-67 and Pro-117). For clarity only FctB-iso residues are labeled. Dashed lines represent hydrogen bonds.

residues, ensuring that it is correctly oriented and close enough to the asparagine carboxamide group for reaction. In our engineered FctB-iso structure, no bond formation occurred until the Val26Phe mutation was made, filling a cavity into which the Lys side chain might otherwise fold and directing it towards the Asn carboxyamide group. Comparison with other isopeptide bond-containing protein domains shows that there is no common pattern to the steric constraints on their bond-forming Lys residues. Their Cγ, Cδ and Cε methylene groups all make van der Waals contact with surrounding hydrophobic residues, but although aromatic residues are commonly involved in a similar way, for example in SpaA[5], RrgB[8,11,12], FimP[14], BP2a[13] and PitB[13], they are far from universal. This is not altogether surprising, however, given the very low levels of sequence identity between the various proteins. The disorder in the catalytic Glu side chain does not seem significant, as the equivalent side chains in other isopeptide bond-containing proteins have a variety of orientations and hydrogen bonding modes with the isopeptide bond moiety. Proximity seems to be sufficient.

The 10 °C increase in melting temperature that accompanies the introduction of the isopeptide bond represents the net gain from replacing the hydrogen bond network that existed in FctB by the covalent crosslink in FctB-iso. It provides a significant increase in thermal stability, and one that would not be sensitive to the redox environment, unlike engineered disulfide bonds. The most useful attribute of the engineered isopeptide bond, however, may be in its resistance to mechanical stress. In FctB-iso, as in other CnaB-type domains[2], it connects the first and last β-strands such that an axial force would be transmitted solely through the isopeptide bonds and the two β-strands they are attached to, leaving the rest of the protein unaffected. Wang *et al.*[31] have shown that this location places the isopeptide bond at a position where the stress concentration is at a maximum[31], and AFM studies have shown experimentally that the protein domains can then withstand forces far in excess of those that rupture other protein domains[22].
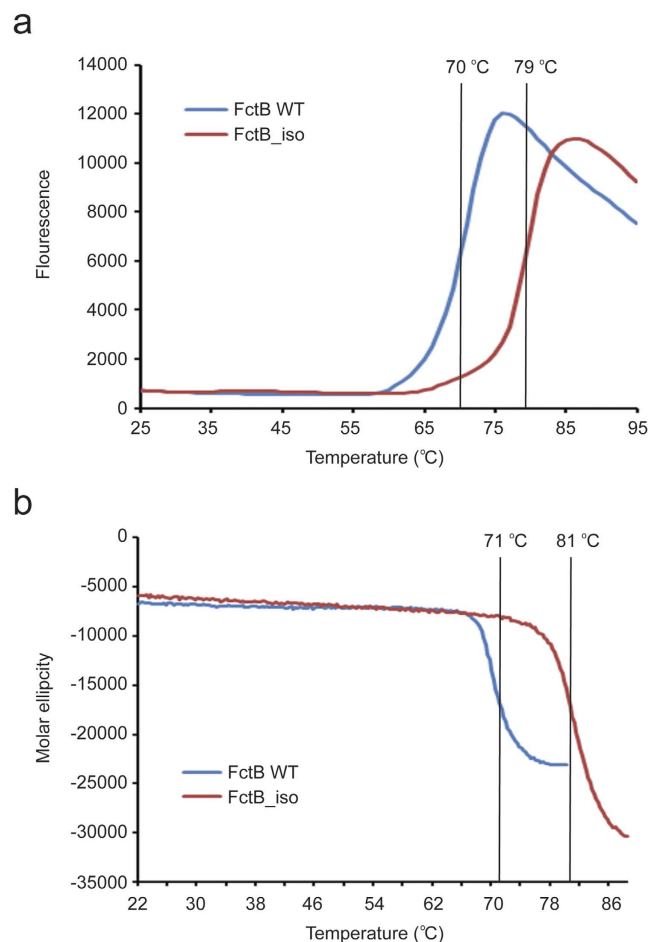
Figure 4. Thermal stability of FctB-iso and WT FctB. Melting curves of FctB-iso (red trace) compared with WT FctB (blue trace) as determined by (**a**) differential scanning fluorimetry (DSF) and (**b**) circular dichroism (CD). The thermal melting temperature of FctB-iso is increased by 9 °C and 10 °C compared to the wild type FctB when measured by DSF and CD, respectively.

In conclusion, we have shown that it is possible, with a small number of mutations (four in this case), to engineer an isopeptide bond crosslink into a protein that lacks any of the requisite residues at the selected site. Bond formation takes place spontaneously, albeit more slowly than in proteins with naturally occurring isopeptide bonds, where they are almost always present when the protein is isolated and are assumed to form during folding. Examples do exist, however, in which there are slow-forming bonds that are apparently facilitated by cooperative interactions[11,15], and further mutations could well increase the speed of bond formation in the present case. What is clear is that the isopeptide bond in FctB-iso adds considerable thermostability. Possible advantages for protein engineering are that an isopeptide bond is both strong and unreactive, and unaffected by redox changes. It is likely that extending isopeptide engineering to protein targets completely unrelated to the Ig-fold will offer much greater challenges. However, as steps towards this goal it should be possible to engineer isopeptide bonds into more divergent Ig-like folds than FctB and potentially into Ig-domains from antibodies.

Methods

Cloning and Mutagenesis.
The gene sequence encoding a C-terminal deletion construct of the extracellular domain of FctB was cloned into the multiple cloning site of vector pProEX HTa (Invitrogen) using 5′ *NarI* and 3′ *EcoRI* restriction-nuclease recognition sites as previously described[32]. The construct encodes the entire extracellular domain minus the N-terminal signal peptide sequence and the 17-residue C-terminal polyproline-II helix tail. The nucleotide and protein sequences for the wild type protein can be found in NCBI GenBank with accession numbers GU250526 and ACZ58644 respectively.

Inverse PCR site-directed mutagenesis was used to modify selected residues[33]. Briefly, a high fidelity DNA polymerase (pfu Ultra II fusion HS, Stratagene) was used for the PCR amplification of the FctB-pProEX HTa construct to give a linearized PCR product with the desired mutation at the 5′ end of the sense primer (5′ phosphorylated primers). Template vector was removed by DpnI digestion, and then re-circularized by intramolecular ligation to produce a modified construct. Ligated DNA was transformed into *Escherichia coli* DH5α cells, and plasmid DNA purified by standard alkaline lysis protocol. Primers used for mutagenesis are listed in Table 2. Mutants were sequence-verified.

Mutation		Sequence
N13K	Fw	/5Phos/GAA**AAA**GTCTTAGAGAGAGCAGGCGATAGTAC
	Rev	/5Phos/GACACTAATGCTAGTTTGAACAGTGC
V26F	Fw	/5Phos/**TTT**GCATTAGAATCAATTGATGC
	Rev	/5Phos/CGAAAATGGGGTACTATCGC
Q67E	Fw	/5Phos/TAT**GAA**AAGCCTTCACAAAATAAAGATTATCAAGC
	Rev	/5Phos/AACACGATAAGTATATTGCCCAACTG
P117N	Fw	/5Phos/AAG**AAC**AAACGGTTAGTAAAACCAATACCGCC
	Rev	/5Phos/AAAAGTAATCGCTGATTTTTCTTCGTCTCC

Table 2. Primer details. Mutated codons are in bold and underlined.

To aid crystallization, N-terminally His$_6$-tagged maltose binding protein (MBP) constructs were made with either a 2-, 3- or 4-residue linker between MBP and FctB. In these constructs the mutated FctB was cloned into pProEX HTa_MBP, an in-house modified pProEX HTa vector series encoding a His$_6$-tagged MBP moiety followed by a variable linker in place of the C-terminal rTEV protease recognition site[34]. The multiple cloning site of pProEX HTa remains intact allowing FctB to be cloned using the same restriction enzyme sites as for the wild type pProEX HTa construct. All mutagenesis and cloning were confirmed by DNA sequencing.

Expression and Purification. All FctB constructs were over-expressed in *E.coli* BL21 (*DE3*) and purified by immobilized metal ion affinity chromatography (IMAC) and size exclusion chromatography (SEC) as described previously[32]. Purification buffers were the same with the exception of the SEC buffer for the MBP-tagged proteins, which contained an additional 5 mM D-maltotriose. Purified recombinant proteins were incubated at room temperature for 7 days in storage buffer (50 mM Tris/HCl pH 7.4, 137 mM NaCl, 2.7 mM KCl) to ensure that any isopeptide bond formation was complete prior to biophysical and crystallization experiments.

Mass Spectrometry Analyses. Accurate molecular masses were determined by infusion ESI-TOF mass spectrometry of purified proteins in 50% acetonitrile and 0.1% formic acid using a Q-STAR XL Hybrid MS/MS system (Applied Biosystems). The raw data were deconvoluted using the Bayesian Protein Reconstruct tool in BioAnalyst software (Applied Biosystem). Confirmation of isopeptide bond formation was provided by trypsin digestion coupled with liquid chromatography-tandem mass spectrometry (LC-MS/MS). Protein samples were subjected to SDS-PAGE analysis to ensure complete homogeneity and the appropriate band was then cleaved from the gel. Gel bands were diced and washed with 50% acetonitrile and 50 mM ammonium bicarbonate to remove staining compounds. Trypsin (10 ng/ul) dissolved in 25 mM ammonium bicarbonate was added to the gel pieces (typically 200 ng of protease) and incubated at 37 °C for 16 h. The supernatant containing digested peptide fragments was collected and analyzed by LC-MS/MS. The MS/MS data were used to search for matches against the wild type FctB sequence using Mascot software (Matrix Science). As the engineered isopeptide bond was designed to link two non-adjacent peptides, those peptides with no matches from the initial Mascot search were selected and analyzed manually to obtain evidence of isopeptide bond formation.

Crystallization and Structure Determination. The isopeptide bond-containing N13K/V26F/Q67E/P117N quadruple mutant of FctB (FctB-iso) was subjected, as a fusion protein with MBP (described above), to initial crystallization experiments by sitting drop vapor diffusion at 18 °C using a Cartesian nanoliter dispensing robot (Genomic systems) and a locally compiled crystallization screen with 480 conditions[35]. Subsequent fine screens were performed by hanging drop vapor diffusion in a 1 + 1 μl format. FctB-iso crystals were grown by mixing 1 μl of protein solution (15 mg/ml, in 50 mM Tris.Cl pH 8.0, 100 mM NaCl, 5 mM D-maltotriose) with 1 μl of precipitant (12.5% PEG 1000, 10% PEG3350, 12.5% MPD, 0.1 M MES/imidazole pH 6.5, 0.02 M sodium formate, 0.02 M ammonium acetate, 0.02 M trisodium citrate, 0.02 M sodium potassium L-tartate and 0.02 M sodium oxamate) at 18 °C. Diffraction data were collected at the Australian Synchrotron on beamline MX-2. Data were processed and scaled with XDS[36] and SCALA[37]. FctB wild type and maltose binding protein structures (PDB codes 3klq and 1anf respectively) were used for molecular replacement with FctB-iso data using PHASER[38]. The structure was refined through cycles of manual building in COOT[39] and maximum likelihood refinement at 2.0 Å resolution using REFMAC[40]. The model was validated using MOLPROBITY[41]. Data collection and refinement statistics are in Table 1. Atomic coordinates and structure factors have been deposited into the Protein Data Bank (www.pdb.org) with accession code 5ttd.

Differential scanning fluorimetry (DSF). Protein thermal unfolding was monitored by the increase in the fluorescence of the fluorophore SyproOrange (Sigma), detected by a real time PCR device (7900HT Fast Real-time PCR System, Applied Biosystems). A total reaction volume of 50 μl contained 5 μl protein solution (30 μM protein in storage buffer), 5 μl 25x Sypro Orange, and 15 μl protein storage buffer. Each experiment was done in triplicate in a 96-well plate with a temperature gradient from 25 °C to 95 °C in steps of 1°/min. Fluorescence emission readings at 600 nm were plotted as a function of temperature. T_m values were fitted to the Boltzmann equation using Microsoft Excel[42].

Circular dichroism (CD).   Wild type FctB and FctB-iso proteins were buffer exchanged into 5 mM sodium phosphate buffer pH 8.0, 50 mM sodium fluoride, to a final concentration of 2.5 μM. CD spectra were recorded on a PiStar-180 (Applied Photophysics) spectrometer. To obtain overall CD spectra, wavelength scans between 180 and 320 nm were collected at 20 °C using a 2 nm bandwidth, 1 nm step size and time per step of 2 sec. The data were collected over 5 accumulations and averaged. To obtain thermal melt curves, data were measured at β-strand minima at 215 nm while the sample temperature was increased from 25 °C to 90 °C at a rate of 1 °C per min.

References

1. Kang, H. J., Coulibaly, F., Clow, F., Proft, T. & Baker, E. N. Stabilizing isopeptide bonds revealed in gram-positive bacterial pilus structure. *Science* **318,** 1625–1628 (2007).
2. Kang, H. J. & Baker, E. N. Intramolecular isopeptide bonds: protein crosslinks built for stress? *Trends in Biochem. Sci.* **36,** 229–236 (2011).
3. Kwon, H., Squire, C. J., Young, P. G. & Baker, E. N. Autocatalytically-generated Thr-Gln ester bond crosslinks stabilize the repetitive Ig-domain shaft of a bacterial cell-surface adhesin. *Proc. Natl. Acad. Sci. USA* **111,** 1367–1372 (2014).
4. Hagan, R. M. *et al.* NMR and theoretical analysis of a spontaneously formed Lys-Asp isopeptide bond. *Angew. Chem. Int. Ed. Engl.* **49,** 8421–8425 (2010).
5. Kang, H. J., Paterson, N. G., Gaspar, A. H., Ton-That, H. & Baker, E. N. The *Corynebacterium diphtheriae* shaft pilin SpaA is built of tandem Ig-like modules with stabilizing isopeptide and disulfide bonds. *Proc. Natl. Acad. Sci. USA* **106,** 16967–16971 (2009).
6. Budzik, J. M. *et al.* Intramolecular amide bonds stabilize pili on the surface of bacilli. *Proc. Natl. Acad. Sci. USA* **106,** 19992–19997 (2009).
7. Pointon, J. A. *et al.* A highly unusual thioester bond in a pilus adhesin is required for efficient host cell interaction. *J. Biol. Chem.* **285,** 33858–33866 (2010).
8. Spraggon, G. *et al.* Supramolecular organization of the repetitive backbone unit of the *Streptococcus pneumoniae* pilus. *PLoS One* **5,** e10919 (2010).
9. Izore, T. *et al.* Structural basis of host cell recognition by the pilus adhesin from *Streptococcus pneumoniae*. *Structure* **18,** 106–115 (2010).
10. Vengadesan, K., Ma, X., Dwivedi, P., Ton-That, H. & Narayana, S. V. L. A model for Group B Streptococcus pilus type 1: the structure of a 35-kDa C-terminal fragment of the major pilin GBS80. *J. Mol. Biol.* **407,** 731–743 (2011).
11. El Mortaji, L. *et al.* The full-length Streptococcus pneumoniae major pilin RrgB crystallizes in a fiber-like structure, which presents the D1 isopeptide bond and provides details on the mechanism of pilus polymerization. *Biochem. J.* **441,** 833–841 (2012).
12. Paterson, N. G. & Baker, E. N. Structure of the full-length major pilin from *Streptococcus pneumoniae*: implications for isopeptide bond formation in Gram-positive bacterial pili. *PLoS One* **6,** e22095 (2011).
13. Nuccitelli, A. *et al.* Structure-based approach to rationally design a chimeric protein for an effective vaccine against Group B Streptococcus infections. *Proc. Natl. Acad. Sci. USA* **108,** 10278–10283 (2011).
14. Persson, K., Esberg, A., Claesson, R. & Stromberg, M. The pilin protein FimP from *Actinomyces oris*: crystal structure and sequence analyses. *PLoS One* **7,** e48364 (2012).
15. Kang, H. J. *et al.* A slow-forming isopeptide bond in the structure of the major pilin SpaD from *Corynebacterium diphtheriae* has implications for pilus assembly. *Acta Crystallogr. D Biol. Crystallogr.* **70,** 968–980 (2014).
16. Shaik, M. M., Maccagni, A., Tourcier, G., Di Guilmi, A. M. & Dessen, A. Structural basis of pilus anchoring by the ancillary pilin RrgC of *Streptococcus pneumoniae*. *J. Biol. Chem.* **289,** 16988–16997 (2014).
17. Forsgren, N., Lamont, R. J. & Persson, K. Two intramolecular isopeptide bonds are identified in the crystal structure of the *Streptococcus gordonii* SspB C-terminal domain. *J. Mol. Biol.* **397,** 740–751 (2010).
18. Larson, M. R. *et al.* Crystal structure of the C-terminal region of *Streptococcus mutans* antigen I/II and characterization of salivary agglutinin adherence domains. *J. Biol. Chem.* **286,** 21657–21666 (2011).
19. Hu, X. *et al.* Autocatalytic intramolecular isopeptide bond formation in Gram-positive bacterial pili: a QM/MM simulation. *J. Amer. Chem. Soc.* **133,** 478–485 (2011).
20. Kang, H. J. & Baker, E. N. Intramolecular isopeptide bonds give thermodynamic and proteolytic stability to the major pilin protein of *Streptococcus pyogenes*. *J. Biol. Chem.* **284,** 20729–20737 (2009).
21. Walden, M., Crow, A., Nelson, M. D. & Banfield, M. J. Intramolecular isopeptide but not internal thioester bonds confer proteolytic and significant thermal stability to the *S. pyogenes* pilus adhesin Spy0125. *Proteins* **82,** 517–527 (2014).
22. Alegre-Cebollada, J., Badilla, C. L. & Fernandez, J. M. Isopeptide bonds block the mechanical extension of pili in pathogenic *Streptococcus pyogenes*. *J. Biol. Chem.* **285,** 11235–11242 (2010).
23. Zakeri, B. & Howarth, M. Spontaneous intermolecular amide bond formation between side chains for irreversible peptide targeting. *J. Amer. Chem. Soc.* **132,** 4526–4527 (2010).
24. Zakeri, B. *et al.* Peptide tag forming a rapid covalent bond to a protein, through engineering a bacterial adhesin. *Proc. Natl. Acad. Sci. USA* **109,** E690–697 (2012).
25. Matsumura, M., Becktel, W., Levitt, M. & Matthews, B. W. Stabilization of phage T4 lysozyme by engineered disulfide bonds. *Proc. Natl. Acad. Sci. USA* **86,** 6562–6566 (1989).
26. Matsumura, M., Signor, G. & Matthews, B. W. Substantial increase of protein stability by multiple disulfide bonds. *Nature* **342,** 291–293 (1989).
27. Linke, C. *et al.* Crystal structure of the minor pilin FctB reveals determinants of Group A streptococcal pilus anchoring. *J Biol Chem* **285,** 20381–20389 (2010).
28. Kang, H. J., Middleditch, M., Proft, T. & Baker, E. N. Isopeptide bonds in bacterial pili and their characterization by X-ray crystallography and mass spectrometry. *Biopolymers* **91,** 1126–1134 (2009).
29. Quiocho, F. A., Spurlino, J. C. & Rodseth, L. E. Extensive features of tight oligosaccharide binding revealed in high-resolution structures of the maltodextrin transport/chemosensory receptor. *Structure* **5,** 997–1015 (1997).
30. Lo, M.-C. *et al.* Evaluation of fluorescence-based thermal shift assays for hit identification in drug discovery. *Analytical Biochemistry* **332,** 153–159 (2004).
31. Wang, B., Xiao, S., Edwards, S. A. & Grate, F. Isopeptide bonds mechanically stabilize Spy0128 in bacterial pili. *Biophys. J.* **104,** 2051–2057 (2013).
32. Linke, C., Young, P. G., Kang, H. J., Proft, T. & Baker, E. N. Purification, crystallization and preliminary crystallographic analysis of the minor pilin FctB from *Streptococcus pyogenes*. *Acta Crystallogr. Sect. F Struct. Biol. Cryst. Commun.* **66,** 177–179 (2010).
33. Ochman, H., Gerber, A. S. & Hartl, D. L. Genetic applications of an inverse polymerase chain reaction. *Genetics* **120,** 621–623 (1988).
34. Ting, Y. T., Batot, G., Baker, E. N. & Young, P. G. Crystallization of a soluble, catalytically-active Type-I signal peptidase from *Staphylococcus aureus*. *Acta Crystallogr. Sect. F Struct. Biol. Commun.* **71,** 61–65 (2015).
35. Moreland, N. *et al.* A flexible and economical medium-throughput strategy for protein production and crystallization. *Acta Crystallogr. Sect. D Biol. Crystallogr.* **61** 1378–1385 (2005).
36. Kabsch, W. Automatic indexing of rotation diffraction patterns. *J. Appl. Crystallogr.* **21,** 67–72 (1988).
37. Evans, P. Scaling and assessment of data quality. *Acta Crystallogr. Sect. D Biol. Crystallogr.* **62,** 72–82 (2006).

38. McCoy, A. J. *et al.* Phaser crystallographic software. *J. Appl. Crystallogr.* **40,** 658–674 (2007).
39. Emsley, P. & Cowtan, K. Coot: model-building tools for molecular graphics. *Acta Crystallogr. Sect. D Biol. Crystallogr.* **60,** 2126–2132 (2004).
40. Murshudov, G. N., Vagin, A. A. & Dodson, E. J. Refinement of macromolecular structures by the maximum-likelihood method. *Acta Crystallogr. Sect. D Biol. Crystallogr.* **53,** 240–255 (1997).
41. Chen, V. B. *et al.* MolProbity: all-atom structure validation for macromolecular crystallography. *Acta Crystallogr. Sect. D Biol. Crystallogr.* **66,** 12–21 (2010).
42. Brown, A. M. A step-by-step guide to non-linear regression analysis of experimental data using a Microsoft Excel spreadsheet. *Computer Methods and Programs in Biomedicine* **65,** 191–200 (2001).
43. Karplus, P. A. & Diederichs, K. Linking crystallographic model and data quality. *Science* **336,** 1030–1033 (2012).
44. Weiss, M. S. & Hilgenfeld, R. On the use of the merging R factor as an indicator of data quality. *J. Appl. Crystallogr.* **30,** 203–205 (1997).

Acknowledgements

Author Contributions

E.N.B., P.G.Y. and C.J.S. conceived and designed the experiments. H.K. and P.G.Y. performed the experiments. P.G.Y., C.J.S and E.N.B. analyzed the data and wrote the manuscript.

Additional Information

Accession codes: The atomic coordinates and structure factors for FctB-iso have been deposited in the Protein Data Bank with the accession codes 5ttd.

Competing financial interests: The authors declare no competing financial interests.

How to cite this article: Kwon, H. *et al.* Engineering a Lys-Asn isopeptide bond into an immunoglobulin-like protein domain enhances its stability. *Sci. Rep.* **7,** 42753; doi: 10.1038/srep42753 (2017).

Publisher's note: Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.